IB/05/50257

PA 1311808

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

April 26, 2005

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE UNDER 35 USC 111.

APPLICATION NUMBER: *60/537,860* ✓
FILING DATE: *January 21, 2004* ✓

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS

P. SWAIN
Certifying Officer

PTO/SB/16 (08-03)
Approved for use through 07/31/2006. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Express Mail Label No. EL0847525608US

INVENTOR(S)

Given Name (first and middle [if any])	Family Name or Surname	Residence (City and either State or Foreign Country)
M. Edward Lisa	Medof Kuttner-Kondo	Pepper Pike, OH Cleveland, OH

Additional inventors are being named on the _____ *separately numbered sheets attached hereto*

TITLE OF THE INVENTION (500 characters max)

Hybrid and Chimeric Polypeptides That Regulate Activation of Complement

Direct all correspondence to: **CORRESPONDENCE ADDRESS**

☒ Customer Number: 24024

OR

☐ Firm or Individual Name					
Address					
Address					
City		State		Zip	
Country		Telephone		Fax	

ENCLOSED APPLICATION PARTS *(check all that apply)*

☒ Specification *Number of Pages* 52 ☐ CD(s), Number _____

☐ Drawing(s) *Number of Sheets* _____ ☐ Other (specify) _____

☐ Application Data Sheet. See 37 CFR 1.76

METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT

☒ Applicant claims small entity status. See 37 CFR 1.27.

☒ A check or money order is enclosed to cover the filing fees.

☒ The Director is herby authorized to charge filing fees or credit any overpayment to Deposit Account Number: 03-0172

☐ Payment by credit card. Form PTO-2038 is attached.

FILING FEE Amount ($): 80.00

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

☐ No.

☐ Yes, the name of the U.S. Government agency and the Government contract number are: _____

Respectfully submitted,

SIGNATURE Kristin J. Frost

TYPED or PRINTED NAME Kristin J. Frost

TELEPHONE 216-622-8895

Date January 21, 2004

REGISTRATION NO. 50,627 *(if appropriate)*

Docket Number: 27708/04053

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

Express Mail Label No. EL0847525608US

INVENTOR(S)		
Given Name (first and middle (if any)	Family Name or Surname	Residence (City and either State or Foreign Country)
M. Edward Lisa	Medof Kuttner-Kondo	Pepper Pike, OH Cleveland, OH

Additional inventors are being named on the _____ separately numbered sheets attached hereto

TITLE OF THE INVENTION (500 characters max)

Hybrid and Chimeric Polypeptides That Regulate Activation of Complement

CERTIFICATE OF EXPRESS MAILING

Date of Deposit: January 21, 2004

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" dididididid under 37 CFR 1.10 on the date indicated above and is addressed to Mail Stop Provisional Patent Appl, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

Robert Watts

Name of Person Signing (Type or Print)

Signed: Robert Watts

27708/04053

HYBRID AND CHIMERIC POLYPEPTIDES THAT REGULATE ACTIVATION OF COMPLEMENT

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

[0001] The present invention was made, at least in part, with support from NIH grant #AI23598. The U.S. government has certain rights in this invention.

BACKGROUND OF THE INVENTION

[0002] The complement system comprises a number of serum proteins that function in the body's immune response to infection and tissue injury. Activation of complement can occur via three pathways, the classical pathway involving the binding of complement component C1q to antigen-antibody complexes, the lectin pathway involving binding of mannose binding lectins to antigens, and the alternative pathway involving binding of complement component C3b to an activator surface such as cell wall polysaccharides of yeast and bacterial microorganisms. Activation of complement results in the formation of anaphylatoxins (C3a and C5a), membrane attack complexes (C5b-9), and opsonins (C3b and C4b) that amplify inflammation and destroy foreign and necrotic cells.

[0003] Complement activation is regulated by a number of plasma and cell associated proteins. Such proteins inactivate specific steps of the classical, lectin, and/or alternative pathway by regulating the activity of C3/C5 convertases or serving as a cofactor for the factor I cleavage of C3b and/or C4b. These proteins are either soluble plasma proteins or membrane proteins (integral or lipid-anchored) expressed on a variety of cell types. These proteins possess many structural similarities.

Decay Accelerating Factor (DAF)

[0004] Decay accelerating factor (DAF, CD55) is a membrane associated regulatory protein that protects self cells from activation of autologous complement on their surfaces. DAF acts by rapidly dissociating C3 and C5 convertases, the central enzymes of the cascade. DAF possesses the most potent decay accelerating activity, and acts on both the classical pathway (C4b2a and C4b2a3b) and alternative pathway (C3bBb and C3bBbC3b) enzymes. DAF, however, does not have cofactor function.

[0005] Structural analyses of DAF have shown that, starting from its N-terminus, it is composed of four ~60 amino acid-long units followed by a heavily O-glycosylated serine (S) and threonine (T) rich stretch, which is, in turn, linked to a posttranslationally-added glycoinositolphospholipid (GPI) anchor. The four 60 amino acid long repeating units are termed complement control protein repeats (CCPs) or short consensus repeats (SCRs). They provide for all of DAF's regulatory activity. The heavily O-glycosylated region serves as a cushion which positions the CCPs at an appropriate distance above the surface membrane. The GPI anchor allows DAF to move freely in the plane of the plasma membrane enabling it to inactivate convertase complexes wherever they assemble.

[0006] The critical role that DAF plays in inhibiting complement activation is evident both from natural disease and studies in animal models employing *Daf* knockout mice. In the human disease paroxysmal nocturnal hemoglobinuria (PNH), mutation in the GPI anchor pathway leading to the absence of DAF renders affected blood cells susceptible to heightened C3b uptake and intravascular hemolysis. In the animal disease models employing the *Daf* knockout, the absence of DAF renders the mice markedly more susceptible to tissue damage in 1) nephrotoxic

serum (NTS) induced nephritis, a model of human membranous glomerulonephritis, 2) dextran sodium sulfate (DSS) induced colitis, a model of inflammatory bowel disease, and 3) anti-acetylcholine receptor (anti-AChR) induced myasthenia gravis, a close model of the human autoimmune disorder.

[0007] Further information regarding the amino acid sequence and physiochemical properties of DAF, and the nucleotide sequence of a cDNA encoding DAF is found in the materials attached hereto. (See "The Complement Facts Book" by Morley and Walport, Academic Press (2000).)

Complement Receptor 1 (CR1)

[0008] Complement receptor 1 (CR1 or the C3b receptor, CD35) is another potent regulator of complement activation. Unlike DAF which functions *intrinsically* to protect the cells that express it, CR1 functions *extrinsically* on targets of complement attack, e.g. pathogens. CR1 is a larger molecule in that, rather than 4 CCPs, it is comprised of 30 CCPs arranged in 4 groups of 7 CCPs termed long homologous repeats (LHRs). Functional analyses have shown that CR1 possesses both decay accelerating activity and cofactor activity for cleavage of C4b and C3b by the serum enzyme, factor I. Early studies showed that among complement regulators, it is the most potent in this latter activity and that it is the only regulator that promotes both initial cleavage of C3b to iC3b and subsequent cleavage of the iC3b intermediate to C3dg, the surface-bound C3b end product.

[0009] Structure-function studies of CR1 have shown that its regulatory activity resides primarily in its three N-terminal LHRs, i.e., LHRs A, B, and C. Functional activity within each 7 CCP LHR is contained essentially in each case in the first 3 CCPs. Recent studies have shown that CR1's potent cofactor activity resides in LHRs B and C, while its decay accelerating activity resides in LHR A.

[0010] Further information regarding the amino acid sequence and physiochemical properties of CR1, and the nucleotide sequence of a cDNA encoding CR1 is found in the materials attached hereto. (See "The Complement Facts Book" by Morley and Walport, Academic Press (2000).)

Membrane Cofactor Protein (MCP)

[0011] MCP (also known as 'CD46') is present on the cell surface of a number of cell types including peripheral blood cells (excluding erythrocytes), cells of epithelial, endothelial and fibroblast lineages, trophoblasts and sperm. MCP has four CCPs and a serine/threonine enriched region in which heavy O-linked glycosylation occurs. MCP also has a transmembrane and cytoplasmic domain. MCP works by binding to the C3b and C4b present on the cell surface thereby targeting C3b and C4b for degradation by factor I, a plasma protease, and thereby destroying any subsequent C3 or C4 convertase activity. Thus, MCP is said to have "cofactor activity". Because MCP is localized on the cell surface, it protects only the cells on which it is present and is therefore said to act in an intrinsic manner. The sequence of a cDNA encoding human MCP has been reported by Lublin et al, J. Exp. Med., (1988) 168:181-194. Further information regarding the amino acid sequence, physiochemical properties, and the nucleotide sequence of a cDNA encoding MCP is found in the materials attached hereto. (See "The Complement Facts Book" by Morley and Walport, Academic Press (2000).)

Effects of Excessive Activation of Complement

[0012] Excessive activation of complement causes damage to normal host tissues in a number of conditions. Some diseases in which complement is known to be activated include systemic lupus erythematosus, acute myocardial infarction, burn, sepsis, stroke and the adult respiratory distress syndrome. Accordingly, it is desirable to have soluble agents that can block complement activation. Such agents would be useful for treating the above-mentioned human diseases and a wide range of other diseases (See Table 1 below).

SUMMARY OF THE INVENTION

[0013] The present invention relates to a family of hybrid and chimeric polypeptides for regulating, more particularly for inhibiting, excessive complement activation.

[0014] The hybrid polypeptides of the present invention comprise at least one functional unit that has been derived from a first complement activation regulatory protein, at least one functional unit that has been derived from a second complement activation regulatory protein,

and a spacer for appropriately separating and spacing the functional units in the hybrid polypeptide from one another. In a preferred embodiment, the hybrid protein (referred to hereafter as a "DAF hybrid") comprises a functional unit, e.g. CCPs 2 and 3 of DAF. Preferably, such functional unit also comprises CCP4 and more preferably CCP1 and CCP4 of DAF. The DAF hybrid protein can also include one or more functional units that have been derived from CR1, e.g., one or more functional units comprising CCPs 8-10 of CR1 or functional units comprising CCPs 15-17 of CR1, or combinations thereof. The DAF hybrid polypeptide can also include one or more functional units that have been derived from MCP, e.g. CCPs 2, 3, and 4. Preferably, the MCP functional unit also comprises CCP1 of MCP. The DAF hybrid protein can also comprise functional units that have been derived from other complement activation regulatory proteins. Examples of such proteins include, but are not limited to, the factor H protein and C4BP. In certain embodiments, the hybrid polypeptide comprises functional units that have been derived from three or more complement activation regulatory proteins, in which each functional unit is separated from the preceding functional unit and following functional unit in the hybrid polypeptide by a spacer.

[0015] The present spacer is a polypeptide that is greater than 200 amino acids in length, preferably greater than 250 amino acids in length. The amino acid sequences of the spacers that are employed in the hybrid proteins of the present invention may be the same or different. The spacer may be a synthetic polypeptide fragment. Alternatively, and preferably, the spacer is derived from a complement activation regulatory protein. In one preferred embodiment, the spacer comprises all or substantially all of CCPs 4-7 of CR1, i.e, amino acid 239 through amino acid 496 of the CR1 sequence shown in the attached document. In another preferred embodiment the spacer comprises all or substantially all of CCPs 11-14 of the CR1 protein. The hybrid proteins of the present invention are based, at least in part, on Applicants discovery that on an equal molar basis, DAF is at least 4 to 5 times more efficient than LHR A of CR1 in inhibiting the classical pathway.

[0016] The chimeric polypeptides of the present invention comprise at least one functional unit that has been derived from a complement activation regulatory protein (referred to hereinafter as the "first functional unit"), a functional unit that has been derived from a protein that is not a complement activation regulatory protein (referred to hereinafter as the "second functional

unit"), and a spacer for separating and appropriately spacing the first functional unit from the second functional unit. The spacer is as described above. The second functional unit can be derived from immunoglobulin (IgG) and may serve to reduce degradation of the chimeric polypeptide following injection into an animal. Alternatively, the second functional unit can be a targeting moiety that enhances binding of the chimeric polypeptide by certain animal tissues. An example of one such targeting moiety is a lipid tail, as shown in attached Figure Z. Such molecule is expected to target the chimeric polypeptide to the membrane bilayer interior, more particularly to areas of translocated acidic phospholipid. (See, Smith, RA (2002) Biochem Soc Trans 30 (Pt6):1037-41.) The second function unit can also be a targeting moiety that enhances binding of the chimeric polypeptide to an implant, or to an extracorporeal surface, e.g.. a hemodialysis membrane. In certain embodiments, the chimeric protein may comprise multiple functional units that have been derived from one or more complement activation regulatory proteins, each of which are separated from one another by a spacer. Thus, the chimeric polypeptide of the present invention can be a hybrid-chimeric polypeptide, e.g. a polypeptide that comprises a functional unit derived from DAF, a functional unit that has been derived from CR1 and a functional unit that has been derived from IgG 4.

[0017] The present invention also provides isolated polynucleotides that encode the hybrid and chimeric polypeptides of the present invention, constructs formed by inserting an isolated polynucleotide of the present invention into an expression vector, and recombinant host cells into which the constructs of the present invention have been incorporated. In addition to the hybrid and/or chimeric polypeptide encoding sequence, such expression vectors comprise regulatory sequences that control or regulate expression of the polypeptide. Examples of suitable host cells are bacterial cells, yeast cells, insect cells, and mammalian cells. The present invention also relates to a process for preparing the hybrid and/or chimeric proteins of the present invention by culturing the cells of the present invention under conditions that promote expression of the hybrid and/or chimeric protein in the cell. For example, the process may be carried out by expressing the hybrid or chimeric protein in Chinese hamster ovary (CHO) cells or COS cells. The hybrid and chimeric proteins of the present invention may then be collected from a cell culture supernatant or cell lysate of the transformed host cells using an affinity column and then eluting the hybrid and/or chimeric protein from the column.

[0018] The present invention also features methods of reducing inflammation characterized by excessive complement activation in an animal subject. In one aspect, the method comprises administering one or more of the present hybrid polypeptides or chimeric polypeptides to an animal subject, particularly a human subject, afflicted with a condition associated with excessive complement activation. Thus, the present invention also relates to methods of treating patients afflicted with any of the diseases listed in Table 1 below. In another aspect, the present method comprises administering an expression vector comprising a polynucleotide that encodes a hybrid polypeptide or a chimeric polypeptide of the present invention to the animal subject.

DETAILED DESCRIPTION OF THE INVENTION

[0019] Described herein are hybrid proteins that comprise at least one functional unit of a first complement activation regulatory protein and at least one functional unit of a second complement activation regulatory protein, particularly a protein that inhibits the activity of C3 and/or C5 convertase. In certain embodiments, such hybrid proteins comprise more than one functional unit from a complement activation regulatory protein. In certain embodiments such hybrid proteins comprises functional units from more than 2 complement activation regulatory proteins. The functional units in the present hybrid proteins are separated and spaced apart by a spacer which is described in greater detail below. The functional units can be located in any order within the hybrid proteins of the present invention.

I. Hybrid Proteins

A. DAF Functional Unit

[0020] In certain embodiments, the present hybrid protein preferably comprises at least one functional unit from DAF. Such functional unit is capable of dissociating C3 and C5 convertases. Thus, the DAF functional unit may comprise CCPs 2 and 3 of DAF, which are sufficient for decay accelerating activity against the classical pathway. Preferably, the DAF functional unit comprise CCPs 2, 3, and 4 of DAF, which are sufficient for decay accelerating activity against both the classical pathway C3 convertase and the alternative pathway C3 convertase. The amino acid sequence of such CCPs may be identical to the native or naturally occurring amino acid sequence of DAF. Alternatively, the amino acid sequence of such CCPs

may be altered slightly, particularly at the amino or carboxy terminus. Such alterations occur when a restriction enzyme site is incorporated into the polynucleotide encoding the CCPs. Such alterations also occur when amino acids are deleted from the N terminus or C terminus of the functional unit. (For example, see Example 1 below in which a number of amino acids are deleted from the C terminus of CCP 4 of DAF.) In certain embodiments the hybrid protein may further comprise CCP 1 of DAF. The amino acid sequence of native DAF is shown in the enclosed attachment. CCP1 extends from and includes amino acid 35 through amino acid 95 of the native DAF protein. CCP2 extends from and includes amino acid through amino acid 97 through amino acid 159 of the native DAF protein. CCP3 extends from and includes amino acid 162-221 of the native DAF protein. CCP 4 extends from and includes 224-284 of the native DAF protein. The functional unit also comprises the amino acids that link CCP1 to CCP2, CCP2 to CCP3, and CCP3 to CCP 4 of the DAF protein.

B. CR1 Functional Units

[0021] The present hybrid protein may also comprise one or more functional units from CR1. Such functional unit is capable of acting as a cofactor for factor I-mediated cleavage of C3b to iC3b and C3f and further cleavage of iC3b to C3c and C3 dg. Such functional unit is also capable of acting as a cofactor for factor I-mediated cleavage of C4b to C4d and Cr4c. (See K. Yazdanbakhsh et al., Blood, 2003). Thus, the hybrid protein of the present invention may comprise substantially all of CCPs 8-10 and/or CCPs 15-17 of CR1. The amino acid sequence of such CCPs may be identical to the native or naturally occurring amino acid sequence of CCPs 8-10 or CCPs 15-17 of CR1. Alternatively, the amino acid sequence of such CCPs may be altered slightly, particularly at the amino or carboxy terminus. Such alterations occur when a restriction enzyme site is incorporated into the polynucleotide encoding the CCPs 8-10 or CCPs 15-17 of CR1 or when amino acids, preferably a few amino acids, are deleted from the amino terminus or the carboxy terminus or both the amino terminus and the carboxy terminus of the functional unit. Preferably, the hybrid protein of the present invention comprises two or more functional units from CR1. Such CR1 functional units may be the same or different. Thus, the hybrid protein of the present invention may comprise two functional units derived from LHR-B or one functional unit derived from LHR B and one functional unit derived from LHR-C of the CR1 protein. CCP 8 of CR1 extends from and includes amino acid 497-556 of the native CR1 protein. CCP 9 of

CR1 extends from and includes amino acids 557-618 of the native CR1 protein. CCP 10 extends from and includes amino acids 619-688 of the native CR1 protein. CCP 15 of CR1 extends from and includes amino acid 947-1006 of the native CR1 protein. CCP 16 of CR1 extends from and includes amino acid 1007-1068 of the native CR1 protein. CCP 17 of CR1 extends from and includes amino acid 1069-1138 of the native CR1 protein.

C. MCP Functional Unit

[0022] The present hybrid protein may also comprise one or more functional units from MCP. Such functional unit is capable of acting as a cofactor for factor I-mediated cleavage of C3b to iC3b and C3f. Thus, the hybrid protein of the present invention may comprise substantially all of CCPs 1-4 of MCP. The amino acid sequence of such CCPs may be identical to the native or naturally-occurring amino acid sequence of CCPs 1-4 of MCP. Alternatively, the amino acid sequence of such CCPs may be altered slightly, particularly at the amino or carboxy terminus. Such alterations occur when a restriction enzyme site is incorporated into the polynucleotide encoding CCPs 1-4 of MCP, or when amino acids are deleted from the amino or carboxy terminus of this functional unit. Preferably, the hybrid protein of the present invention comprises two or more functional units from MCP, or alternatively, a functional unit from MCP and a functional unit from CR1. CCP 1 of MCP extends from and includes amino acid 35-95 of the native MCP protein. CCP 2 of MCP extends from and includes amino acid 96-158 of the native MCP protein. CCP3 extends from and includes amino acid 159-224 of the native MCP protein. CCP4 extends from and includes amino acid 225-285 of the native MCP protein.

D. Spacer

[0023] The hybrid proteins of the present invention comprise one or more spacers. Each spacer in the present hybrid and/or chimeric proteins separate and appropriately space the functional units of the present hybrid and/or chimeric proteins from one another. Such spacer is a polypeptide that is greater than 200 amino acids in length, preferably greater than 250 amino acids in length. The amino acid sequences of the spacers that are employed in the hybrid proteins of the present invention may be the same or different. In one preferred embodiment, the spacer comprises all or substantially all of CCPs 4-7 of CR1, i.e., amino acid 239 through amino

acid 496 of the CR1 sequence shown in the attached document. In another preferred embodiment the spacer comprises all or substantially all of CCPs 11-14 of the CR1 protein. As used herein the term substantially all means that the spacer may lack a few, e.g. 1-10 amino acids from the N terminal and/or the C terminal of the spacer. The spacer may also comprise some amino acids that result from incorporating a restriction enzyme site into the spacer. Thus, the spacer may comprise a few amino acids at the N terminus or C terminus that are different from the amino acids that are found at the N terminus or C terminus of the CCP4-7 fragment that is derived from native CR1 or the CCP11-14 fragment that is derived from native CR1.

[0024] Optionally, the hybrid proteins of the present invention may further include a tag, i.e., a second protein or one or more amino acids, preferably from about 2 to 65 amino acids, that are added to the amino or carboxy terminus of the hybrid protein. Typically, such additions are made to stabilize the protein or to simplify purification of an expressed recombinant form of the hybrid protein Such tags are known in the art. Representative examples of such tags include sequences which encode a series of histidine residues, the epitope tag FLAG, the Herpes simplex glycoprotein D, beta-galactosidase, maltose binding protein, or glutathione S-transferase.

[0025] The present invention also encompasses hybrid protein proteins in which one or more amino acids are altered by post-translation processes or synthetic methods. Examples of such modifications include, but are not limited to, glycosylation, iodination, myristoylation, and pegylation.

II. Chimeric Proteins.

[0026] The chimeric proteins of the present invention comprise one or more functional units of complement activation regulatory protein as described above and one or more functional units derived from a protein that is not a complement activation regulatory protein. Examples of functional units that are not derived from complement activation regulatory proteins include functional units that are derived from an immunoglobulin, particularly IgG4, and that serve to reduce degradation of the chimeric polypeptide following injection into an animal. Thus, the chimeric protein may include the hinge, CH2, and CH3 domains of IgG4. Alternatively, the second functional unit can be a targeting moiety that enhances binding of the chimeric

polypeptide by certain animal tissues. An example of such targeting moiety is a lipid tail, as shown in attached Figure Z. In certain embodiments, the chimeric protein may comprise multiple functional units that have been derived from one or more complement activation regulatory proteins, each of which are separated from one another by a spacer. Thus, the chimeric polypeptides of the present invention can be a hybrid, chimeric polypeptide.

[0027] Preparation of the Hybrid and Chimeric Proteins

[0028] The present hybrid proteins and chimeric proteins of the present invention are prepared using polynucleotides that encode such proteins and expression systems.

[0029] The functional units and spacers employed in the present hybrid and/or chimeric proteins can be made by obtaining total (t) or messenger (m) RNA from an appropriate tissue, cell line or white blood cells. Suitable RNA (total or messenger) is also available commercially. Blood can be drawn from a human subject and peripheral blood mononuclear cells (PBMCs) can be purified by Ficoll-Paque density centrifugation. Total RNA from PBMCs should contain CR1 and IgG4. Cell lines can be grown in controlled climate incubators with appropriate cell culture media. DAF and MCP are fairly ubiquitous proteins. Thus, these proteins can be found in most cell lines, e.g., the HeLa cell line.

[0030] Following isolation of suitable RNA, the RNA is reverse transcribed to cDNA using commerically available reagents and standard protocols, e.g., the Superscript protocol of Invitrogen. Once the appropriate cDNA is made, polymerase chain reaction (PCR) can be used in conjunction with DNA polymerases and oligonucleotide primer pairs (20 to 30 nucleotides in length) to amplify DAF, MCP, CR1 and/or IgG4 cDNA. One primer will be at the 5' end of the cDNA (for example at the start codon ATG or, in the case of the constant heavy chain, further downstream at the start of the constant heavy region 1 [CH1] and one primer will be at the 3' end of the cDNA, e.g. at the stop codon TAG, TAA, or TGA). The PCR products are then subcloned into vectors such as pT7Blue (pT7B) (Novagen, Madison, WI) and sequenced to confirm that the correct cDNA was amplified.

[0031] Expression Systems For Producing The Hybrid Proteins

[0032] The present hybrid proteins can be produced in procaryotic and eucaryotic cells each using different expression vectors that are appropriate for each host cell. Eucaryotic expression system such as the baculoviral or mammalian cells are described below.

[0033] The following are examples of expression vectors which may be used for gene expression in an eucaryotic expression system. The plasmid, pMSG, uses the promoter from mouse mammary tumor virus long terminal repeat (MMTV). Suitable host cells for pMSG expression are chinese hamster ovary (CHO) cells, HeLa cells and mouse Lkt negative cells (Lee, F., et al., 1981, Nature 294:228-232). The vector, pSVL, uses the SV40 late promoter. For high transient expression, suitable host cells for pSVL expression are COS cells (Sprague, J. et al., 1983, J. Virol. 45:773-781). The vector, pRSV, uses Rous Sarcoma Virus promoter. Suitable host cells for pRSV expression are mouse fibroblast cells, lymphoblastoid cells and COS cells (Gorman, Padmanabhan and Howard, 1983, Science 221:551-553).

[0034] Baculovirus expression vectors can also be used. These vectors are stably expressed in insect cells such as sf9 (Luckow, V. A. and Summers, M. D., 1988, Bio/Technology 6:47-55; Miller, L. K., 1988, Ann. Re. Microbiology 42:177-199).

[0035] Hybrid proteins of the invention can also be produced in a procaryotic expression system. The following are examples of expression vectors which can be expressed in procaryotic expression systems. The pOX expression series using the oxygen-dependent promoter can be expressed in E. coli. (Khosla, G., et al., 1990, Bio/Technology 8:554-558). pRL vector which uses the strong pL promoter of lambda phage (Reed, R. R., 1981, Cell 25:713-719; Mott, J. D., et al., 1985, Proc. Natl. Acad. Sci. U.S.A. 82:88-92) and the pKK223-3 vector which uses a hybrid promoter derived from the fusion between the promoters of the tryptophan and lactose operons of E. coli. (Borsius, J. and Holy, A., 1984, Proc. Natl. Acad. Sci. U.S.A. 81:6929-6933) can be used for expression in E. coli.

[0036] Suitable vectors for yeast expression are also well known in the art, e.g., Sleep, D., Belfield, D. P. and Goodey, A. R., 1990, Bio/Technology 8:42-46; Sakai, A. et al., 1991,

Bio/Technology 9:1382-1385; Kotula, L. and Curtis, P. J., 1991, Bio/Technology 9:1386-1389, all of which are herein incorporated by reference.

[0037] Production, Quantitation, Purification And Analysis Of The Hybrid Proteins.

[0038] Once a recombinant cell line that expresses the hybrid protein has been isolated, the secreted proteins are identified and verified with regard to their predicted structure. Various methods can be used to identify and characterize the expressed hybrid proteins. The presence of secreted hybrid proteins can be verified by immunoprecipitation with monoclonal antibodies to one or the other fragment, e.g., antibodies that bind to CCP's, 2, 3, or 4 of DAF or to LHR B or C of CR1.

[0039] Another method that could be used with present hybrid and/or chimeric proteins is a double immunoprecipitation, using two monoclonal antibodies of different specificities in succession. Pre-clearance of culture supernatant with one antibody would result in a negative immunoprecipitation with the second antibody. This method would verify that a single protein expresses both CR1 and DAF epitopes.

[0040] Alternatively, the hybrid DAF-CR1 protein, can be identified by Western blot. For example, after SDS-PAGE and transfer to nitrocellulose, blots can be developed with either anti-CR1 antibodies or anti-DAF monoclonal antibodies. The expressed bispecific recombinant protein would be reactive with both antibodies, again demonstrating the presence of specific DAF and CR1 epitopes in the hybrid protein.

[0041] Identification of the present hybrid and/or chimeric proteins can also be accomplished by ELISA. For example, a rabbit polyclonal antibody specific for either LHR B or C of CR1 or CCP's 2, 3, or 4 of DAF can be used to coat plastic microtiter ELISA plates, followed by the addition of culture supernatant from the recombinant cell line expressing the DAF-CR1 hybrid and incubation with the capture polyclonal antibody. A monoclonal anti-DAF or anti-CR1 second antibody, the specificity of which is different from the capture antibody, can be subsequently used. A positive reaction would indicate the presence of both epitopes on the hybrid or chimeric protein.

[0042] An ELISA can also be used to quantitate the levels of the DAF-CR1 hybrid protein in culture supernatants or any other unpurified solutions containing the chimeric protein by comparison to standard curve of known quantities of purified DAF-CR1 hybrid protein. Quantitation of DAF-CR1 hybrid protein would be useful for determination of production rates in recombinant cell lines, determination of protein concentration in partially purified preparations, and for determination of protein concentration in plasma for in vivo experiments.

[0043] The hybrid and/or chimeric proteins of the present invention can be purified from recombinant cell culture supernatant by a variety of standard chromatographic procedures, including but not limited to immunoaffinity chromatography, ion exchange chromatography, gel filtration chromatography, reverse-phase high pressure liquid chromatography (HPLC), lectin affinity chromatography, or chromatofocusing. For example, small quantities of culture supernatant containing serum supplement can be purified using immunoaffinity chromatography with, e.g., anti-CR1 or anti-DAF monoclonal antibodies. DAF-CR1 hybrid protein bound to the immobilized antibody can be eluted in purified form by use of a chaotropic solution.

[0044] Once the hybrid and/or chimeric protein is purified, its amino acid sequence can be deduced by direct protein sequence analysis using an automated system. The presence of N- and O-linked carbohydrates can be determined by use of specific endoglycosidase enzymes (Chavira, R. et al., 1984, Anal. Biochem. 136:446). Further characterizations of its biochemical structure can also be performed, including but not limited to pI determination by isoelectric focusing, hydrophilicity analysis, X-ray crystallographic analysis, and computer modeling.

Functional Characterization Of The Present Hybrid And/Or Chimeric Proteins

[0045] The hybrid proteins of the present invention have the ability to function both as a cofactor for Factor I and as a decay accelerating factor. In vitro assays can be performed to measure these biological activities (Medof, M. et al., 1984, J. Exp. Med. 160:1558; Masaki, T. et al., 1992, J. Biochem 111:573). As described in the examples, assays for cofactor activity and for decay accelerating activity are used to demonstrate both these complement regulatory functions for the present hybrid protein. The consequence of either cofactor or decay accelerating activity, or in the case of a DAF-CR1 or DAF-MCP hybrid protein, both activities in combination, is the

inactivation of C3/C5 convertases. Another suitable in vitro assay demonstrates that the present hybrid protein is capable of inhibiting C5 convertase activity as measured by the production of C5a (Moran, P. et al., 1992, J. Immunol. 149:1736, herein incorporated by reference). Additional assays, as described in the examples below, demonstrate that the present hybrid proteins inhibit the complement-induced lysis of cells via the classical and alternative pathways.

Demonstration Of In Vivo Therapeutic Activity Of The Present Hybrid And/Or Chimeric Proteins

[0046] The Arthus reaction is an inflammatory response caused by the interaction of antigen in tissue with circulating antibody. It has been used as a classic example of a localized in vivo inflammatory response, and is characterized by the formation of immune complexes, complement activation, inflammatory cell recruitment, edema and tissue damage (Bailey, P. and Sturm, A., 1983, Biochem. Pharm 32:475). Experimentally, a reversed passive Arthus reaction can be established in an animal model by i.v. injection with antigen and subsequent challenge with antibody. Using guinea pigs as an animal model, the in vivo therapeutic efficacy of the hybrid and/or chimeric proteins of the invention can be evaluated.

[0047] Additional animal models with relevance to various clinical human diseases can also be used to test the in vivo efficacy of complement activation blockers. These include, but are not limited to: myocardial ischemia/reperfusion injury (acute myocardial infarction; Weisman, H. F. et al., 1990, Science 249:146); cerebral ischemic injury (stroke; Chang, L. et al., 1992, J. Cerebr. Blood Flow Metab. 12:1030); lung injury (ARDS; Hosea, S. et al., 1980, J. Clin. Invest. 66:375); xenograft rejection (transplants; Leventhal, J. et al., 1993, Transplantation 55:857); burn injury (Caldwell, F, et al., 1993, J. Burn Care Rehab. 14:420); acute pancreatitis (Steer, M. 1992, Yale J. Biol. Med. 65:421), nephritis (Pichler, R. et al., 1994, Am. J. Pathol. 144:915), cardiopulmonary bypass (Nilsson, L. et al., 1990, Artif. Organs 14:46), and multiple sclerosis (Linington, C. et al., 1989, Brain 112:895).

Administration of the Present Hybrid and Chimeric Proteins to Animal Subjects

[0048] The present hybrid and chimeric proteins can be combined with an appropriate pharmaceutical formulation and administered to an animal subject, particularly a human subject, by a variety of routes, including, but not limited to, intravenous bolus injection, intravenous

infusion, intraperitoneal, intradermal, intramuscular, subcutaneous, and intranasal routes. The administration of the present hybrid proteins in vivo will enable the protein to bind endogenous C3/C5 convertases and inhibit the generation of additional C3b and C5b, of C3a and C5a anaphylatoxins, and of C5b-9 lytic complexes. The complement regulatory activities of the present hybrid proteins can therefore function to inhibit in vivo complement activation and the inflammatory sequelae that accompany it, such as neutrophil recruitment and activation, autolysis of host cells, and edema. The present hybrid and/or chimeric proteins can be used for the therapy of diseases or conditions that are mediated by inordinate and/or excessive activation of the complement system. These include, but are not limited to: tissue damage due to ischemia-reperfusion following myocardial infarction, aneurysm, stroke, hemorrhagic shock, or crush injury; burns; endotoxemia and septic shock; adult respiratory distress syndrome (ARDS); hyperacute rejection of grafts; cardiopulmonary bypass and pancreatitis. Autoimmune disorders including, but not limited to, systemic lupus erythematosis, rheumatoid arthritis, and multiple sclerosis, can also be treated with the hybrid and/or chimeric proteins of the invention (also see Table 1).

TABLE 1

Potential Clinical Targets of Protein of the Invention	
Alternative Pathway	Classical Pathway
Reperfusion injury	Autoimmune diseases
Cerebral infarction (stroke)	
	Systemic lupus erythematosus
Acute myocardial infarction	
	Rheumatoid arthritis
Hypovolemic shock	Glomerulonephritis
Multiple organ failure	
	Hemolytic anemia
Crush injury	Myasthenia gravis
Intestinal ischemia	Multiple sclerosis
	Vasculitis
	Inflammatory bowel diseases
	Bullous diseases
	Urticaria
	Paroxysmal nocturnal hemoglobinuria
	Cryoglobulinemia
Inflammatory disorders	
	Inflammatory disorders
Adult respiratory distress syndrome	
	Septic shock & endotoxemia
Thermal injury (burn & frostbite)	
Post-pump syndrome (cardiopulmonary bypass & hemodialysis)	
Crohn's disease	
Sickle cell anemia	
pancreatitis	
Adverse drug reactions	
	Transplant rejection
Drug allergy	Hyperacute allograft
Radiographic contrast media allergy	
IL-2 induced vascular leakage syndrome	
Transplant rejection	
Hyperacute xenograft	

[0049] Various delivery systems are known and can be used to deliver the hybrid and/or chimeric proteins of the invention, such as encapsulation in liposomes, or controlled release devices. The hybrid and/or chimeric proteins of the invention can also be administered extracorporeally, e.g., pre-conditioning donor organs prior to transplantation. The hybrid and/or chimeric proteins of the invention can be formulated in a pharmaceutical excipient in the range of approximately 10 μg/kg and 10 mg/kg body weight for in vivo or ex vivo treatment.

Administration of Polynucleotides That Encode the Present Hybrid and/or Chimeric Proteins to an Animal Subject

[0050] The present invention also relates to therapeutic methods in which polynucleotides that encode and express the present hybrid and/or chimeric polypeptides are introduced into a subject in need of the same, i.e. a subject, particularly a human subject, with a disorder associated with increased complement activation. Polynucleotides encoding and expressing one or more hybrid and/or chimeric polypeptide can be introduced into cells of the subject using any of a variety of methods known in the art to achieve transfer of DNA molecules into cells. For example, DNA encoding and expressing the hybrid and/or chimeric polypeptide can be incorporated into liposomes and targeted to and internalized by the cells of the subject. Polynucleotides encoding the hybrid and/or chimeric polypeptide can also be incorporated into plasmids that are introduced into cells of the subject by transfection. The hybrid and/or chimeric polypeptide encoding polynucleotides can also be introduced into cells using viruses. Such viral "vectors" can have DNA or RNA genomes. Numerous such viral vectors are well known to those skilled in the art. Viral vectors that have polynucleotide sequences encoding a DAF-CR1 hybrid polypeptide, for example, cloned into their genomes are referred to as "recombinant" viruses. Transfer of DNA molecules using viruses is particularly useful for transferring polynucleotide sequences into particular cells or tissues of an animal. Such techniques are commonly known in the art as gene therapy.

[0051] Expression vectors normally contain sequences that facilitate gene expression. An expression vehicle can comprise a transcriptional unit comprising an assembly of a protein encoding sequence and elements that regulate transcription and translation. Transcriptional regulatory elements generally include those elements that initiate transcription. Types of such elements include promoters and enhancers. Promoters may be constitutive, inducible or tissue specific. Transcriptional regulatory elements also include those that terminate transcription or provide the signal for processing of the 3' end of an RNA (signals for polyadenylation). Translational regulatory sequences are normally part of the protein encoding sequences and include translational start codons and translational termination codons. There may be additional sequences that are part of the protein encoding region, such as those sequences that direct a protein to the cellular membrane, a signal sequence for example.

[0052] The hybrid and/or chimeric polypeptide encoding polynucleotides that are introduced into cells are preferably expressed at a high level (i.e., the introduced polynucleotide sequence produces a high quantity of the hybrid and/or chimeric polypeptide within the cells) after introduction into the cells. Techniques for causing a high-level of expression of polynucleotide sequences introduced into cells are well known in the art. Such techniques frequently involve, but are not limited to, increasing the transcription of the polynucleotide sequence, once it has been introduced into cells. Such techniques frequently involve the use of transcriptional promoters that cause transcription of the introduced polynucleotide sequences to be initiated at a high rate. A variety of such promoters exist and are well known in the art. Frequently, such promoters are derived from viruses. Such promoters can result in efficient transcription of polynucleotide sequences in a variety of cell types. Such promoters can be constitutive (e.g., CMV enhancer/promoter from human cytomegalovirus) or inducible (e.g., MMTV enhancer/promoter from mouse mammary tumor virus). A variety of constitutive and inducible promoters and enhancers are known in the art. Other promoters that result in transcription of polynucleotide sequences in specific cell types, so-called "tissue-specific promoters," can also be used. A variety of promoters that are expressed in specific tissues exist and are known in the art. For example, promoters whose expression is specific to neural, liver, epithelial and other cells exist and are well known in the art. Methods for making such DNA molecules (i.e., recombinant DNA methods) are well known to those skilled in the art.

[0053] In the art, vectors refer to nucleic acid molecules capable of mediating introduction of another nucleic acid or polynucleotide sequence to which it has been linked into a cell. One type of preferred vector is an episome, i.e., a nucleic acid capable of extrachromosomal replication. Other types of vectors become part of the genome of the cell into which they are introduced. Vectors capable of directing the expression of inserted DNA sequences are referred to as "expression vectors" and may include plasmids, viruses, or other types of molecules known in the art.

[0054] Typically, vectors contain one or more restriction endonuclease recognition sites which permit insertion of the hybrid polypeptide encoding sequence. The vector may further comprise a marker gene, such as for example, a dominant antibiotic resistance gene, which encode compounds that serve to identify and separate transformed cells from non-transformed cells.

[0055] One type of vector that can be used in the present invention is selected from viral vectors. Viral vectors are recombinant viruses which are generally based on various viral families comprising poxviruses, herpesviruses, adenoviruses, parvoviruses and retroviruses. Such recombinant viruses generally comprise an exogenous polynucleotide sequence (herein, a polynucleotide encoding the hybrid and/or chimeric polypeptide) under control of a promoter which is able to cause expression of the exogenous polynucleotide sequence in vector-infected host cells.

[0056] One type of viral vector is a defective adenovirus which has the exogenous polynucleotide sequence inserted into its genome. The term "defective adenovirus" refers to an adenovirus incapable of autonomously replicating in the target cell. Generally, the genome of the defective adenovirus lacks the sequences necessary for the replication of the virus in the infected cell. Such sequences are partially or, preferably, completely removed from the genome. To be able to infect target cells, the defective virus contains sufficient sequences from the original genome to permit encapsulation of the viral particles during *in vitro* preparation of the construct. Other sequences that the virus contains are any such sequences that are said to be genetically required "in *cis*."

[0057] Preferably, the adenovirus is of a serotype which is not pathogenic for man. Such serotypes include type 2 and 5 adenoviruses (Ad 2 or Ad 5). In the case of the Ad 5 adenoviruses, the sequences necessary for the replication are the E1A and E1B regions. Methods for preparing adenovirus vectors are described in U.S. Patent No. 5,932,210, U.S Patent No. 5,985,846, and U.S. Patent No. 6,033,908.

[0058] More preferably, the virus vector is an immunologically inert adenovirus. As used herein the term "immunologically inert" means the viral vector does not encode viral proteins that activate cellular and humoral host immune responses. Methods for preparing immunologically inert adenoviruses are described in Parks et al., *Proc Natl Acad Sci USA* 1996; 93(24) 13565-70; Leiber, A. et al., *J. Virol.* 1996; 70(12) 8944-60; Hardy s., et al, *J. Virol.* 1997, 71(3): 1842-9; and Morsy et al, *Proc. Natl. Acad. Sci. USA* 1998. 95: 7866-71, all of which are specifically incorporated herein by reference. Such methods involve Cre-*loxP* recombination. In vitro, Cre-*loxP* recombination is particularly adaptable to preparation of recombinant adenovirus and offers

a method for removing unwanted viral nucleotide sequences. Replication deficient recombinant adenovirus lacks the E1 coding sequences necessary for viral replication. This function is provided by 293 cells, a human embryonic kidney cell line transformed by adenovirus. First generation adenoviruses are generated by co-transfecting 293 cells with a helper virus and a shuttle plasmid containing the foreign gene of interest. This results in the packaging of virus that replicates both the foreign gene and numerous viral proteins. More recently, 293 cells expressing Cre recombinase, and helper virus containing essential viral sequences and with a packaging signal flanked by *loxP* sites, have been developed (See Parks et al.) In this system, the helper virus supplies all of the necessary signals for replication and packaging *in trans*, but is not packaged due to excision of essential sequences flanked by *loxP*. When 293-Cre cells are co-transfected with this helper virus, and a shuttle plasmid (pRP1001) containing the packaging signal, nonsense "filler DNA", and the foreign gene, only an adenovirus containing filler DNA and the foreign gene is packaged (LoxAv). This results in a viral recombinant that retains the ability to infect target cells and synthesize the foreign gene, but does not produce viral proteins.

[0059] Another type of viral vector is a defective retrovirus which has the exogenous polynucleotide sequence inserted into its genome. Such recombinant retroviruses are well known in the art. Recombinant retroviruses for use in the present invention are preferably free of contaminating helper virus. Helper viruses are viruses that are not replication defective and sometimes arise during the packaging of the recombinant retrovirus.

[0060] Non-defective or replication competent viral vectors can also be used. Such vectors retain sequences necessary for replication of the virus. Other types of vectors are plasmid vectors.

[0061] The methods also involve introduction of polynucleotides encoding the present hybrid and/or chimeric polypeptides into an animal subject in the context of cells (e.g., *ex vivo* gene therapy).

EXAMPLES

[0062] The following examples contained herein are intended to illustrate but not limit the invention.

EXAMPLE 1: Hybrid Protein 313

[0063] A hybrid protein, 313, comprising a decay accelerating functional unit derived from DAF, a cofactor 1 functional unit derived from CR1, and a spacer comprised of CCPs 4-7 of CR1 was made by recombinant techniques. The DAF portion of the hybrid protein 3I3 was constructed using *DAF13.2.l/pBTKS* and two primers *DSIGEB* and *DAF3P* in a PCR reaction (Vent polymerase [New England Biolabs] with the following times: 94°C 3min [initial melting]; 94°C 1min, 55°C 1min15sec, 72°C, 1min15sec for 25 cycles; and 72°C 7min [final extension]). *DSIGEB* is a 42 nucleotide ("nt") primer that has the sequence 5'-ATA TAC GAA TTC AGA TCT ATG ACC GTC GCG CGG CCG AGC GTG-3'. *DAF3P* is a 35nt primer that has the sequence 5'-ACA GTG CTC GAG CAT TCA GGT GGT GGG CCA CTC CA-3'. The resultant PCR product was named *DAF1*. It contained DAF's signal sequence followed by CCPs 1, 2, 3 and 4 ending with cysteine 249 (Cys-249) in CCP4. Upstream of the signal sequence, two restriction enzyme sites were built in, BglII (A▼GATCT) and 5' of BglII, EcoRI (G▼AATTC). Three prime (3') of CCP4 and encompassing part of the Cys-249 codon (TGC), the restriction enzyme site XhoI (C▼TCGAG) was inserted. DAF1 was subcloned into *pT7B* and fully sequenced.

[0064] The CR1 portion of the hybrid protein 3I3 was constructed using *CR1/AprM8*. *CR1/AprM8* was cut with the restriction enzyme NsiI (ATGCA▼T) releasing several pieces, two of which were recovered (1094nts and 1350nts) and subcloned into *pGEM7Zf(+)*. The "1094" fragment (encompassing nts 557 to 1670 of CR1) was PCRed using the primers *CR1094X*(5') and *CR1094N*(3'). *CR1094X* is a 41nt primer having the sequence 5'-ATA TAC CTC GAG TCC TAA CAA ATG CAC GCC TCC AAA TGT GG-3'. It has an XhoI site. *CR1094N* is a 34nt primer having the sequence 5'-ACA GTG ATG CAT TGG TTT GGG TTT TCA ACT TGG C-3'. It has an NsiI site. This set of primers produces a sequence from the linker between CCP3 and CCP4 of CR1 into CCP8 of CR1. PCR conditions were the same as those for DAF1.

The "1350" fragment, encompassing nts 1671 to 3020 of CR1, was PCRed using primers *CR1350N*(5') and *CR1B3P*(3'). *CR1350N* is a 41nt primer having the sequence 5'-ATA TAC ATG CAT CTG ACT TTC CCA TTG GGA CAT CTT TAA AG-3'. It has an NsiI site. *CR1B3P* is a 57nt primer having the sequence 5'-ACA GTG AGA TCT TTA GTG ATG GTG ATG GTG ATG AAT TCC ACA GCG AGG GGC AGG GCT-3'. It has a BglII site. PCR conditions were the same as those for *DAF1* except the 25 cycle extension time at 72° was 2min, not 1min15sec. This set of primers produces a sequence from CCP8 of CR1 to the end of CCP14 (in LHRB, specifically, ...SSPAPRCGI) with a C-terminal 6XHis tag and stop codon. These PCR fragments were subcloned into *pT7B*.

[0065] It is noteworthy that the natural linker between CCP3 and CCP4 of CR1 is "IIPNK". Due to the insertion of the XhoI restriction site, the hybrid protein's linker between DAF CCP4 and CR1 CCP4 is "SSPNK".

[0066] Two colonies of each CR1 PCR/*pT7B* subcloning product were sent for sequencing: *CR828XN2*, *CR828XN3*, *CR1300NBC* and *CR1300NBF*. Data obtained for *CR828XN3* and *CR1300NBF* indicated that these clones had correct nucleotide sequences.

[0067] The vector pSG5 (Stratagene) was cut with the restriction enzymes EcoRI and BglII to accommodate the insertion of *DAF1* (EcoRI to XhoI), *CR828XN3*(XhoI to NsiI) and *CR1300NBF* (NsiI to BglII). The vector and the three fragments were ligated using Promega T4 DNA ligase, and transformed into DH5α maximum efficiency competent cells. Agarose gel electrophoresis showed that 5 of 8 colonies contained the vector and insert. The cDNA from colony 3 was used for transfection into COS cells using Lipofectamine (Invitrogen) reagent. The supernatant was harvested two days later. Western blots using 2H6, an anti-DAF CCP4 antibody, and an anti-His tag antibody indicated the presence of the hybrid protein.

[0068] The hybrid protein 3I3 was tested in a 1) classical pathway C3 convertase hemolytic assay along with a recombinant soluble DAF protein composed of CCPs 1, 2, 3 and 4 for comparison; and 2) normal human serum (NHS) hemolytic assay along with the recombinant DAF protein and recombinant soluble CR1 (sCR1) protein for comparison. 3I3 performed as

well as DAF in the classical pathway assay and at least as efficient as sCR1 (many times better than the recombinant DAF protein) in the NHS assay.

EXAMPLE 2: Hybrid Protein DAFBB

[0069] Another hybrid protein, ,DAFBB, which comprises a decay accelerating functional unit derived from DAF, two cofactor 1 functional units derived from CR1, a spacer comprised of CCPs 4-7 separating the functional unit of DAF from the first cofactor 1 functional unit of CR1, and a second spacer CCPs 11-14 of CR1 separating the first cofactor 1 functional unit of CR1 from the second cofactor 1 functional unit of of CR1 was prepared by recombinant techniques. More specifically, DAFBB was prepared by adding an additional LHRB of CR1 to 3I3 To add the additional cofactor LHR, 3I3 was cut with BamHI and a BamHI fragment (nucleotide #1861 to 3210) from CR1 in AprM8 was introduced. The BamHI fragment could enter the plasmid in either the correct or reverse orientation. Screening with SmaI found several clones with the correct nucleotide orientation.

EXAMPLE 3: Chimeric Protein Mini

[0070] To increase half life of the hybrid protein, while minimizing complement activation, as a starting point, part of the constant heavy region of IgG4 was PCRed and connected 3' to nucleotides encoding a decay accelerating functional unit derived from DAF, and a spacer derived from CR1. (see "DAF-CR1 Hybrid/Chimeric Designs".) The resulting protein is referred to hereafter as "Mini". Mini is composed of DAF CCPs1,2,3,4 + CR1 CCP4,5,6,7 (part of LHRA) + IgG4, last amino acid (valine) of CH1- Hinge-CH2-CH3. The domains of IgG4 were PCR'd from pHC-huCg4, a gift of Gary McLean, 2222 Health Sciences Mall, Vancouver, B.C., Canada. The primers used in this PCR reaction were:

IgG45: 5'-ATA TAC GAA TTC TGG TTG AGT CCA AAT ATG GTC CC-3' and

IgG43: 5'-ACA GTG AGA TCT TTA TCA TTT ACC CGG AGA CAG GGA G-3'.

[0071] Per 100µl PCR reaction, 2U Vent polymerase (New England Biolabs), 71.5ng pHC-Cg4 (7157 bp), 50 pmol of each primer, and 10mM of each dNTP were used. PCR settings were 1

initial denaturing cycle of 94°C 3 min; 25 cycles of 94°C 1 min, 55°C 1 min 15 sec, and 72°C 1 min; and a final extension cycle of 72°C 7 min. A 700 bp fragment was recovered with the QIAquick gel purification kit. The fragment "IgG4 PCR" was ligated into the pT7Blue (pT7B) blunt vector (Novagen). Nova Blue (Novagen) and XL1Blue (Stratagene) competent cells were transformed with the ligation mixture and plated on Ampicillin 50, Tetracycline 15, IPTG/Xgal LB agar plates. Of four colonies screened, 3 had inserts. All insert-containing plasmids were sequenced. IgG41/pT7Blue ("IgG4/pT7B 1", 5-18-03) cDNA was good. It was subsequently cut with EcoRI ("E") and BglII ("B") in Promega Buffer H ("H") and the 700bp band was purified with QIAquick. IgG41 E/B and pSG5 E/B/H were ligated using the Quick Ligation method (New England Biolabs) and transformed into DH5αmax (Invitrogen) competent cells. Of four colonies selected, three cDNAs had inserts when cut with EcoRI and BglII in Buffer H (Promega).

[0072] IgG41/pSG5-1 was cut with EcoRI (~4800bp linearized), purified, shrimp alkaline phosphatase ("SAP")-treated, purified again, and quick ligated to 3I3A/EcoRI-cut fragment (gel purified, 1-23-03, 1650bp) which contains the DAF and CR1 portions of the sequence. DH5αmax competent cells were transformed with the ligation mix. To screen resulting colonies, cutting the cDNA with BglII will yield two bands of 4100bp and 2350bp if the orientation is correct; while it will yield 5650bp and 700bp bands if incorrect. Colony I ("Mini") had the correct band pattern. Mini I-1 was maxi-prepped and cDNA was checked (uncut and BglII-cut). The cDNA was transfected into COS cells. Protein was visualized on a 5% SDS-PAGE gel with the monoclonal antibody IA10 against CCP1 of DAF (See Western Blot).

[0073] Note that the IgG45 primer codes for a slightly different link between CR1 CCP7 and the Hinge of IgG4 (in Mini) than the link between CR1 CCP7 and MCP CCP1 (in Micki). It results in "GILV" ("V" is the last amino acid of the CH1 domain of IgG4) instead of "GILGH" which is found in Micki and is also the normal link between CR1 CCP7 and CCP8 (and is therefore what is found in 3I3 and DAFBB hybrids).

EXAMPLE 4 Hybrid Protein Micki

[0074] A hybrid protein, referred to hereafter as Micki, comprising a decay accelerating functional of DAF, a cofactor 1 functional unit derived from MCP, and a spacer derived from CR1 was prepared. MCP cDNA (with 3'-end sequence encoding GPI-anchor addition) in PEE14 was used. More MCP cDNA in DH5α (Wizard SV mini-prep) ("MCP-GPI (A)") was subsequently prepared. Primers for the MCP portion of Micki (DAF CCPs 1,2,3,4- CR1 CCPs 4,5,6,7- MCP CCPs 1,2,3,4+ 2 amino acids (VS) of MCP STP region + 6xHis) are:

MCP5: 5'-ATA TAC GAA TTC TGG GTC ACT GTG AGG AGC CAC CAA CAT TTG AAG C-3'; and

MCP3: 5'-ACA GTG AGA TCT TTA GTG ATG GTG ATG GTG ATG CGA CAC TTT AAG ACA CTT TGG AAC-3'.

[0075] The PCR reaction used Vent polymerase from New England Biolabs. The MCP PCR fragment was cut with EcoRI ("E") and BglII ("B") in Promega Buffer H ("H"). The Quick Ligase method (New England Biolabs) was used to ligate E/B/H-cut MCP PCR and E/B/H-cut pSG5. DH5α maximum efficiency competent cells were transformed with the mixture. Four colonies of MCP/pSG5 were picked, the cDNA was extracted from the bacteria and cut with E/B/H. All had an insert. Colonies 1 and 2 were sent for sequencing (MCP/pSG5/1 and MCP/pSG5/2. Nucleotide 680 was found to be incorrect, a G when it should be an A. This changes the amino acid at that position. Primers were made to do site-directed mutagenesis to correct the error:

G680A5: 5'-GCA CGA TTT ATT GTG GTG ACA ATT CAG TGT GGA GTC GTG C-3'; and

G680A3: 5'-GCA CGA CTC CAC ACT GAA TTG TCA CCA CAA TAA ATC GTG C-3'.

[0076] Stratagene's Quick Change II Kit was used. Cycle conditions were 1 cycle at 95°C 30sec; and 12 cycles at 95°C 30sec, 55°C 1min, and 68°C 5min. The reaction mixture was cut

with DpnI and a small quantity was used in transforming XL1Blue supercompetent cells. cDNA from two colonies, MCP1A and MCP1B, was sequenced. MCP1A sequence was correct.

[0077] MCP1A/pSG5 was cut with EcoRI in Buffer H and purified with Qiagen PCR purification kit. The purification product was treated with shrimp alkaline phosphatase ("SAP"), and purified again. Quick ligation method was used to ligate EcoRI- and SAP-treated MCP1A/pSG5 and EcoRI-treated and gel purified 1650 base pair piece from 3I3A which adds the DAF and CR1 portions of the cDNA. After transformation, 8 resulting colonies were mini-prepped. cDNA was cut with BglII and some cDNA with EcoRI independently as well. Correct pieces would be 4.1kB/2.45kB (BglII) and 4.9kB/1.65kB (EcoRI). Note that EcoRI-cut only shows an insert is present, not whether the orientation is correct. An incorrect orientation of insert would give 5.75kB/0.8kB fragments (BglII). Colonies 5 and 7 looked good.

[0078] Micki 5A cDNA was maxi-prepped (Qiagen). TAE agarose gel was run the next day to show uncut, BglII-cut, and EcoRI-cut cDNA. Bands were as expected. This cDNA was used for COS cell transfection. Micki 5A protein was visualized by western blot with monoclonal antibodies IA10 (against DAF CCP1) and GB24 (against MCP CCPs 3 and 4). Size was between 90kDa and 100kDa.

[0079] Testing of the Hybrid Proteins

[0080] Hemolytic assays show that the activity of DAFBB>3I3≥DAF. DAFBB appears to outperform sCR1 in normal human serum and classical pathway C5 convertase hemolytic assays. Additionally, DAFBB has been shown to have cofactor activity in a factor I/C3b assay.

[0081] Hemolytic Assays

[0082] Hemolytic assays are performed to assess the activity of the components of the classical and alternative pathways of complement or the activity of the regulators of complement activation (RCA) proteins. Classical pathway activity is assessed using antibody-sensitized sheep erythrocytes (EshA) and can be undertaken in a variety of ways. Whole serum can be used, or purified components of the complement cascade can be used in a classical pathway C3 or C5 convertase hemolytic assay.

[0083] Experiments were undertaken using normal human, pig and rat serum to assess animal models for study. Human DAF has been found to be active against convertases formed in pig serum (J.M. Perez de la Lastra *et al.* 2000, J. Immunol. 165:2563). It is active at high concentrations against convertases formed in rat serum (C.L. Harris *et al.*, 2000, Immunology 100:462). In contrast, human CR1 is highly active against convertases formed in rat serum. Classical pathway C3 convertase and C5 convertase assays utilizing purified human complement components were also done to compare the hybrids' performance to soluble DAF and soluble CR1.

Whole Serum Experiments

[0084] Serum (human, pig or rat) was titrated to a Z score of approximately 1, defined as an average of 1 lesion per sheep cell (A.P. Gee, 1983, "Molecular Titration of Components of the Classical Complement Pathway" in Methods in Enzymology 93:339). To each tube was added $1x10^7$ EshA, an RCA protein and the serum in $DGVB^{++}$ (dextrose gelatin veronal buffer with calcium and magnesium) to a total volume of 200ul. This mixture was shaken for 30min in a 37°C water bath. GVBE (gelatin veronal buffer with EDTA) was subsequently added to stop the hemolytic reaction. The extent of lysis was determined by reading the ED_{412} of the hemoglobin red supernatant.

Classical Pathway C3 Convertase Hemolytic Assay

[0085] EshA ($1x10^7$ cells/tube) were sequentially shaken and spun down with 30SFU (site-forming units) of human C1 (ART) for 15min, 15SFU of human C4 (Quidel) for 20min, and sufficient human C2 (ART) for a Z score of 1 for 5min, all at 30°C. Following formation of the classical pathway C3 convertase C4b2a, regulators were added for 15min at 30°C to assess their relative decay-accelerating activity. Guinea pig serum in GVBE (1:40 dilution) was subsequently added for 1hr at 37°C to allow formation of the terminal membrane complex. The cells were spun down and the OD_{412} of the hemoglobin red supernatant was read to determine the extent of cell lysis.

Classical Pathway C5 Convertase Hemolytic Assay

[0086] The classical pathway C5 convertase hemolytic assay was run over a 2-day period. EshA ($1x10^7$ cells/tube) were shaken sequentially with 60SFU of human C1 (ART) for 15min, 60SFU of human C4 (Quidel) for 20min, and 10SFU human C2 (calculated after decay in a C3 convertase assay) (ART) along with 15SFU of human C3 (gift of C. Mold) for 5min, all in a 30°C water bath, to form the classical pathway C5 convertase C4b3b2a. Following this loading of complement components, 200ul DGVB++ was added to each tube and the tubes were shaken for 2hr in a 30°C water bath to decay the C2. The cells were spun down, resuspended in DGVB++, and decay continued overnight with the cells at 4°C. The following day, the cells were shaken with 60SFU of human C1 for 15min in a 30°C water bath, followed by sufficient human C2 for a Z score of 1 (5min, 30°C). RCA proteins were added for 15min at 30°C to accelerate the decay of the C5 convertase. Subsequently, human C5 (Quidel) (1:250 dilution) was added for 5 min at 30°C, then guinea pig C6-9 in DGVBE (dextrose gelatin veronal buffer with EDTA) (1:150 dilution) for 1hr at 37°C. Cells were spun down and the OD_{412} of the hemoglobin red supernatant was read to determine the extent of cell lysis.

Cofactor Experiments

[0087] The RCA proteins CR1, factor H and membrane cofactor protein MCP can act as cofactors for the factor I cleavage of C3b to smaller fragments (reviewed in M. Botto, "C3" [p. 88] in The Complement FactsBook, ed. B.J. Morley and M.J. Walport, 2000, San Diego: Academic Press). The cofactor activity of CR1 resides in its LHRs B and C (S.C. Makrides *et al.*, 1992, J. Biol. Chem. 267:24754; K.R. Kalli *et al.*, 1991, J. Exp. Med. 174:1451; M. Krych *et al.*, 1994, J. Biol. Chem. 269:13273). All three regulators can allow factor I to cleave C3b to iC3b. Factor I with CR1 can additionally cleave to iC3b to C3c. When the disulfide bridges are reduced, an SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis) gel will show the following bands:

C3b: α' (115kDa), β (75kDa)

iC3b: α'-68, α'-43, β (75kDa), and a small soluble fragment C3f

C3c: α'-27, α'-40, β (75kDa), and "membrane-attached" C3dg (40kDa)

(A. Sahu *et al.*, 1998, J. Immunol. 160:5596; A.M. Rosengard *et al.*, 2002 PNAS 99:8808).

[0088] To confirm that the hybrid retained the cofactor activity of CR1, human C3b (ART Lot 20P, 20ng), was mixed with human factor I (ART Lot 6P, 60ng). Factor H (500ng) or DAFBB (20ng) or a control of 10mM PO4 buffer with 145mM NaCl, pH 7.3 was added for a total volume of 10ul and incubated for 19hr in a 37° water bath.

THE COMPLEMENT *FactsBook*

Bernard J. Morley
Mark J. Walport
Imperial College School of Medicine
Hammersmith Campus, London, UK

ACADEMIC PRESS
A Harcourt Science and Technology Company
San Diego San Francisco New York Boston
London Sydney Tokyo

31

This book is printed on acid-free paper.

Academic Press
A division of Harcourt Science and Technology Company
24–28 Oval Road, London NW1 7DX, UK
http://www.hbuk.co.uk/ap/

Academic Press
525 B Street, Suite 1900, San Diego, California 92101-4495, USA
http://www.apnet.com

ISBN 0-12-733360-6

Library of Congress Catalog Card Number: 99-65744

A catalogue for this book is available from the British Library

Typeset by Mackreth Media Services, Hemel Hempstead, UK
Printed in Great Britain by Redwood Books, Trowbridge, Wiltshire

00 01 02 03 04 RB 9 8 7 6 5 4 3 2 1

Lloyd B. Klickstein, Brigham and Women's Hospital, Boston, MA, USA.
Joann M. Moulds, University of Texas Medical School, Houston, TX, USA

Other names

Complement receptor type 1, C3b/C4b receptor, CD35, immune adherence receptor.

Physicochemical properties

CR1 is a type 1 integral membrane glycoprotein of 2044 amino acids[1-4], of which the leader sequence comprises either 41 or 46 amino acids; there are two possible translation initiation sites. The C-terminal transmembrane region contains 25 amino acids and there are 43 residues in the C-terminal cytoplasmic domain. The N-terminal residue is blocked[5,10], compatible with pyrollidone amide cyclization or N-terminal alkylation of Gln47.

There are at least four major structural allotypes described in humans[6-9], the most common form is CR1*1 (F or A), and all further descriptions will focus on that human form except where specifically noted. The other forms are CR1*2 (B or S), CR1*3 (Cor F') and CR1*4 (D).

pI[9] 7.1

M_r (K) 205–250 (depending on cell source and electrophoresis system[5-13]).

Allotype	Approx. M_r (reduced)	Approx. M_r (unreduced)
CR1*1	220–250	190–210
CR1*2	250–280	220–250
CR1*3	190–220	160–190
CR1*4	>280	>250

CR1 from polymorphonuclear leukocytes migrates M_r (K) 5 larger than that from erythrocytes due to altered *N*-linked glycosylation[11].

N-linked glycosylation sites 25 (61, 161, 257, 320, 415, 452, 514, 583, 707, 770, 865, 902, 964, 1033, 1157, 1220, 1315, 1486, 1509, 1539, 1545, 1610, 1673, 1768, 1913)

N-linked glycosylation contributes approximately 20–25 K to the molecular weight of the CR1*1[7,11-13]. Protein sequence data from erythrocyte CR1[3] supports occupancy of sites at 514 and 964. Similarly, sites at 320 and 770 are unoccupied. Occupancy of the other sites is unknown. There is no detectable O-linked glycosylation[12].

Structure

CR1 has an extracellular region comprised of a linear array of 30 CCP units of 59–75 residues each[2-5]. There are 120 cysteine residues and all are believed involved in disulfide links, based on structural homology to β_2 glycoprotein 1[14]. An extended linear structure has been confirmed by electron microscopy[15]. The N-terminal 28 CCPs are further organized as four tandem, long homologous repeats of seven CCP units each[2,4]. The predicted transmembrane region was confirmed by deletion mutagenesis, which resulted in a soluble form of the protein[15,16].

Function

CR1 has long been recognized as the receptor for C3b and C4b fragments, and recently as a receptor for C1q[17]. CR1 also binds iC3b, but relatively poorly[18]. Human erythrocyte CR1 mediates binding of complement-opsonized immune complexes or microorganisms to the cell[19]. These bound complexes or particles are then carried through the bloodstream to the spleen or liver where they are removed[20-24]. CR1 on neutrophils and monocytes can mediate phagocytosis if the cells are primed or activated[25-27]. CR1 on B cells and dendritic cells participates in localization of antigen for presentation to T cells[28-32]. CR1 on all cell types is a cofactor for factor I-mediated cleavage of C3b to iC3b and C3f, and further cleavage of iC3b to C3c and C3d,g. CR1 is a cofactor for factor I-mediated cleavage of C4b to C4c and C4d. CR1 also accelerates the otherwise spontaneous decay of the C3 and C5 convertases of the classical pathway (C4b2a and C4b2a3b) as well as that of the corresponding alternative pathway convertases (C3bBb and C3bBbC3b)[1,2,33]. These activities may be either intrinsic or extrinsic (located on the same surface as the CR1 or not)[34,35].

Tissue distribution

CR1 as a type 1 transmembrane protein is found on all erythrocytes, B cells, polymorphonuclear leukocytes, monocytes, follicular dendritic cells and glomerular podocytes and is also found on a subset of T cells[2,34,37]. CR1 is absent on NK cells[38]. A soluble form is found in serum at a concentration reported at 30–60 ng/ml[39,40], however this is an overestimate as the monoclonal antibodies used have repeated epitopes in CR1[4,41].

Regulation of expression

CR1 is constitutively expressed on the previously mentioned cells. It is slowly lost from the surface of erythrocytes over the normal life of the cells. This loss is greatly accelerated in patients with immune complex diseases such as systemic lupus erythematosus[42-44] and is an acquired phenomenon, not an hereditary predisposition to illness[45]. Ninety per cent of neutrophil CR1 is intracellular[46-48], located in secretory vesicles distinct from azurophilic or specific granules[49]. Upon neutrophil activation with chemotactic peptides or other stimuli, this intracellular CR1 is mobilized to the cell surface[46,47].

Protein sequence[3,4,50]

```
MCLGRMGASS PRSPEPVGPP APGLPFCCGG SLLAVVVLLA LFVAWGQCNA   50
PEWLPFARPT NLTDEFEFPI GTYLNYECRP GYSGRPFSII CLKNSVWTGA  100
KDRCRRKSCR NPPDPVNGMV HVIKGIQFGS QIKYSCTKGY RLIGSSSATC  150
IISGDTVIWD NETPICDRIP CGLPPTITNG DFISTNRENF HYGSVVTYRC  200
NPGSGGRKVF ELVGEPSIYC TSNDDQVGIW SGPAPQCIIP NKCTPPNVEN  250
GILVSDNRSL FSLNEVVEFR CQPGFVMKGP RRVKCQALNK WEPELPSCSR  300
VCQPPPDVLH AERTQRDKDN FSPGQEVFYS CEPGYDLRGA ASMRCTPQGD  350
WSPAAPTCEV KSCDDFMGQL LNGRVLFPVN LQLGAKVDFV CDEGFQLKGS  400
SASYCVLAGM ESLWNSSVPV CEQIFCPSPP VIPNGRHTGK PLEVFPPGKA  450
VNYTCDPHPD RGTSFDLIGE STIRCTSDPQ GNGVWSSPAP RCGILGHCQA  500
PDHFLFAKLK TQTNASDFPI GTSLKYECRP EYYGRPFSIT CLDNLVWSSP  550
KDVCKRKSCK TPPDPVNGMV HVITDIQVGS RINYSCTTGH RLIGHSSAEC  600
ILSGNAAHWS TKPPICQRIP CGLPPTIANG DFISTNRENF HYGSVVTYRC  650
NPGSGGRKVF ELVGEPSIYC TSNDDQVGIW SGPAPQCIIP NKCTPPNVEN  700
GILVSDNRSL FSLNEVVEFR CQPGFVMKGP RRVKCQALNK WEPELPSCSR  750
VCQPPPDVLH AERTQRDKDN FSPGQEVFYS CEPGYDLRGA ASMRCTPQGD  800
WSPAAPTCEV KSCDDFMGQL LNGRVLFPVN LQLGAKVDFV CDEGFQLKGS  850
SASYCVLAGM ESLWNSSVPV CEQIFCPSPP VIPNGRHTGK PLEVFPPGKA  900
VNYTCDPHPD RGTSFDLIGE STIRCTSDPQ GNGVWSSPAP RCGILGHCQA  950
PDHFLFAKLK TQTNASDFPI GTSLKYECRP EYYGRPFSIT CLDNLVWSSP 1000
KDVCKRKSCK TPPDPVNGMV HVITDIQVGS RINYSCTTGH RLIGHSSAEC 1050
ILSGNTAHWS TKPPICQRIP CGLPPTIANG DFISTNRENF HYGSVVTYRC 1100
NLGSRGRKVF ELVGEPSIYC TSNDDQVGIW SGPAPQCIIP NKCTPPNVEN 1150
GILVSDNRSL FSLNEVVEFR CQPGFVMKGP RRVKCQALNK WEPELPSCSR 1200
VCQPPPEILH GEHTPSHQDN FSPGQEVFYS CEPGYDLRGA ASLHCTPQGD 1250
WSPEAPRCAV KSCDDPLGQL PHGRVLFPLN LQLGAKVSFV CDEGFRLKGS 1300
SVSHCVLVGM RSLWNNSVPV CEHIFCPNPP AILNGRHTGT PSGDIPYGKE 1350
ISYTCDPHPD RGMTFNLIGE STIRCTSDPH GNGVWSSPAP RCELSVRAGH 1400
CKTPEQFPFA SPTIPINDFE FPVGTSLNYE CRPGYFGKMF SISCLENLVW 1450
SSVEDNCRRK SCGPPPEPFN GMVHINTDTQ FGSTVNYSCN EGFRLIGSPS 1500
TTCLVSGNNV TWDKKAPICE IISCEPPPTI SNGDFYSNNR TSFHNGTVVT 1550
YQCHTGPDGE QLFELVGERS IYCTSKDDQV GVWSSPPPRC ISTNKCTAPE 1600
VENAIRVPGN RSFFSLTEII RFRCQPGFVM VGSHTVQCQT NGRWGPKLPH 1650
CSRVCQPPPE ILHGEHTLSH QDNFSPGQEV FYSCEPSYDL RGAASLHCTP 1700
QGDWSPEAPR CTVKSCDDFL GQLPHGRVLL PLNLQLGAKV SFVCDEGFRL 1750
KGRSASHCVL AGMKALWNSS VPVCEQIFCP NPPAILNGRH TGTPFGDIPY 1800
GKEISYACDT HPDRGMTFNL IGESSIRCTS DPQGNGVWSS PAPRCELSVP 1850
AACPHPPKIQ NGHYIGGHVS LYLPGMTISY TCDPGYLLVG KGFIFCTDQG 1900
IWSQLDHYCK EVNCSFPLFM NGISKELEMR KVYHYGDYVT LKCEDGYTLE 1950
GSPWSQCQAD DRWDPPLAKC TSRAHDALIV GTLSGTIFPI LLIIFLSWII 2000
LKHRKGNNAH ENPKEVAIHL HSQGGSSVHP RTLQTNEENS RVLP
```

The leader sequence is underlined and the potential *N*-linked glycosylation sites are indicated (N).

Protein modules[3,4,50]

1 or 6–46	Leader peptide	exon 1
47–106	CCP1, begin LHR-A	exon 2
107–168	CCP2	exon 3/4
169–238	CCP3	exon 5
239–300	CCP4	exon 5
301–360	CCP5	exon 6
361–423	CCP6	exon 7/8
424–496	CCP7, end LHR-A	exon 9
497–556	CCP8, begin LHR-B	exon 10
557–618	CCP9	exon 11/12
619–688	CCP10	exon 13
689–750	CCP11	exon 13
751–810	CCP12	exon 14
811–873	CCP13	exon 15/16
874–946	CCP14, end LHR-B	exon 17
947–1006	CCP15, begin LHR-C	exon 18
1007–1068	CCP16	exon 19/20
1069–1138	CCP17	exon 21
1139–1200	CCP18	exon 21
1201–1260	CCP19	exon 22
1261–1323	CCP20	exon 23/24
1324–1399	CCP21, end LHR-C	exon 25
1400–1459	CCP22, begin LHR-D	exon 26
1460–1521	CCP23	exon 27/28
1522–1591	CCP24	exon 29
1592–1653	CCP25	exon 29
1654–1713	CCP26	exon 30
1714–1776	CCP27	exon 31/32
1777–1851	CCP28, end LHR-D	exon 33
1852–1911	CCP29	exon 34
1912–1972	CCP30	exon 35
1977–2001	Transmembrane region	exon 36/37
2002–2044	Cytoplasmic region	exon 38

The ligand-binding sites are[4,16,51–55]:

47–300	CCPs 1–4	C4b-binding site (lower affinity for C3b)
497–750	CCPs 8–11	C3b-binding site (lower affinity for C4b)
947–1200	CCPs 15–18	C3b-binding site (lower affinity for C4b)
1400–1851	CCPs 22–28	A C1q-binding site

Chromosomal location

Human[56,57]: 1q32.
Telomere ... MCP ... CR1 ... CR2 ... DAF ... C4bp ... Centromere
Factor H (Cfh) maps to 1q32 but has not been physically linked with other members of the RCA.
Mouse[58,59]: chromosome 1q, 40 cM.
Telomere ... Crry ... CR1/CR2 ... Cfh ... C4bp ... Centromere

cDNA sequence[3,4,59,60]

```
TTTTGTCCCG GAACCCCGCA GCCCTCCCCA CACTCTGGGC GCGGAGCACA ATGATTGGTC   60
ACTCCTATTT TCGCTGAGCT TTTCCTCTTA TTTCAGTTTT CTTCGAGATC AAATCTGGTT  120
TGTAGATGTG CTTGGGGAGA ATGGGGGCCT CTTCTCCAAG AAGCCCGGAG CCTGTCGGGC  180
CGCCGGCGCC CGGTCTCCCC TTCTGCTGCG GAGGATCCCT GCTGGCGGTT GTGGTGCTGC  240
TTGCGCTGCC GGTGGCCTGG GGTCAATGCA ATGCCCCAGA ATGGCTTCCA TTTGCCAGGC  300
CTACCAACCT AACTGATGAG TTTGAGTTTC CCATTGGGAC ATATCTGAAC TATGAATGCC  360
GCCCTGGTTA TTCCGGAAGA CCGTTTTCTA TCATCTGCCT AAAAAACTCA GTCTGGACTG  420
GTGCTAAGGA CAGGTGCAGA CGTAAATCAT GTCGTAATCC TCCAGATCCT GTGAATGGCA  480
TGGTGCATGT GATCAAAGGC ATCCAGTTCG GATCCCAAAT TAAATATTCT TGTACTAAAG  540
GATACCGACT CATTGGTTCC TCGTCTGCCA CATGCATCAT CTCAGGTGAT ACTGTCATTT  600
GGGATAATGA AACACCTATT TGTGACAGAA TTCCTTGTGG GCTACCCCCC ACCATCACCA  660
ATGGAGATTT CATTAGCACC AACAGAGAGA ATTTTCACTA TGGATCAGTG GTGACCTACC  720
GCTGCAATCC TGGAAGCGGA GGGAGAAAGG TGTTTGAGCT TGTGGGTGAG CCCTCCATAT  780
ACTGCACCAG CAATGACGAT CAAGTGGGCA TCTGGAGCGG CCCCGCCCCT CAGTGCATTA  840
TACCTAACAA ATGCACGCCT CCAAATGTGG AAAATGGAAT ATTGGTATCT GACAACAGAA  900
GCTTATTTTC CTTAAATGAA GTTGTGGAGT TTAGGTGTCA GCCTGGCTTT GTCATGAAAG  960
GACCCCGCCG TGTGAAGTGC CAGGCCCTGA ACAAATGGGA GCCGGAGCTA CCAAGCTGCT 1020
CCAGGGTATG TCAGCCACCT CCAGATGTCC TGCATGCTGA GCGTACCCAA AGGGACAAGG 1080
ACAACTTTTC ACCTGGGCAG GAAGTGTTCT ACAGCTGTGA GCCCGGCTAC GACCTCAGAG 1140
GGGCTGCGTC TATGCGCTGC ACACCCCAGG GAGACTGGAG CCCTGCAGCC CCCACATGTG 1200
AAGTGAAATC CTGTGATGAC TTCATGGGCC AACTTCTTAA TGGCCGTGTG CTATTTCCAG 1260
TAAATCTCCA GCTTGGAGCA AAAGTGGATT TTGTTTGTGA TGAAGGATTT CAATPAAAAG 1320
GCAGCTCTGC TAGTTACTGT GTCTTGGCTG GAATGGAAAG CCTTTGGAAT AGCAGTGTTC 1380
CAGTGTGTGA ACAAATCTTT TGTCCAAGTC CTCCAGTTAT TCCTAATGGG AGACACACAG 1440
GAAAACCTCT GGAAGTCTTT CCCTTTGGAA AAGCAGTAAA TTACACATGC GACCCCCACC 1500
CAGACAGAGG GACGAGCTTC GACCTCATTG GAGAGAGCAC CATCCGCTGC ACAAGTGACC 1560
CTCAAGGGAA TGGGGTTTGG AGCAGCCCTG CCCCTCGCTG TGGAATTCTG GGTCACTGTC 1620
AAGCCCCAGA TCATTTTCTG TTTGCCAAGT TGAAAACCCA AACCAATGCA TCTGACTTTC 1680
CCATTGGGAC ATCTTTAAAG TACGAATGCC GTCCTGAGTA CTACGGGAGG CCATTCTCTA 1740
TCACATGTCT AGATAACCTG GTCTGGTCAA GTCCCAAAGA TGTCTGTAAA CGTAAATCAT 1800
GTAAAACTCC TCCAGATCCA GTGAATGGCA TGGTGCATGT GATCACAGAC ATCCAGGTTG 1860
GATCCAGAAT CAACTATTCT TGTACTACAG GGCACCGACT CATTGGTCAC TCATCTGCTG 1920
AATGTATCCT CTCGGGCAAT GCTGCCCATT GGAGCACGAA GCCGCCAATT TGTCAACGAA 1980
TTCCTTGTGG GCTACCCCCC ACCATCGCCA ATGGAGATTT CATTAGCACC AACAGAGAGA 2040
ATTTTCACTA TGGATCAGTG GTGACCTACC GCTGCAATCC TGGAAGCGGA GGGAGAAAGG 2100
TGTTTGAGCT TGTGGGTGAG CCCTCCATAT ACTGCACCAG CAATGACGAT CAAGTGGGCA 2160
TCTGGAGCGG CCCGGCCCCT CAGTGCATTA TACCTAACAA ATGCACGCCT CCAAATGTGG 2220
AAAATGGAAT ATTGGTATCT GACAACAGAA GCTTATTTTC CTTAAATGAA GTTGTGGAGT 2280
TTAGGTGTCA GCCTGGCTTT GTCATGAAAG GACCCCGCCG TGTGAAGTGC CAGGCCCTGA 2340
ACAAATGGGA GCCGGAGCTA CCAAGCTGCT CCAGGGTATG TCAGCCACCT CCAGATGTCC 2400
TGCATGCTGA GCGTACCCAA AGGGACAAGG ACAACTTTTC ACCCGGGCAG GAAGTGTTCT 2460
ACAGCTGTGA GCCCGGCTAT GACCTCAGAG GGGCTGCGTC TATGCGCTGC ACACCCCAGG 2520
GAGACTGGAG CCCTGCAGCC CCCACATGTG AAGTGAAATC CTGTGATGAC TTCATGGGCC 2580
AACTTCTTAA TGGCCGTGTG CTATTTCCAG TAAATCTCCA GCTTGGAGCA AAAGTGGATT 2640
TTGTTTGTGA TGAAGGATTT CAATTAAAAG GCAGCTCTGC TAGTTATTGT GTCTTGGCTG 2700
GAATGGAAAG CCTTTGGAAT AGCAGTGTTC CAGTGTGTGA ACAAATCTTT TGTCCAAGTC 2760
CTCCAGTTAT TCCTAATGGG AGACACACAG GAAAACCTCT GGAAGTCTTT CCCTTTGGAA 2820
AAGCAGTAAA TTACACATGC GACCCCCACC CAGACAGAGG GACGAGCTTC GACCTCATTG 2880
GAGAGAGCAC CATCCGCTGC ACAAGTGACC CTCAAGGGAA TGGGGTTTGG AGCAGCCCTG 2940
CCCCTCGCTG TGGAATTCTG GGTCACTGTC AAGCCCCAGA TCATTTTCTG TTTGCCAAGT 3000
TGAAAACCCA AACCAATGCA TCTGACTTTC CCATTGGGAC ATCTTTAAAG TACGAATGCC 3060
GTCCTGAGTA CTACGGGAGG CCATTCTCTA TCACATGTCT AGATAACCTG GTCTGGTCAA 3120
GTCCCAAAGA TGTCTGTAAA CGTAAATCAT GTAAAACTCC TCCAGATCCA GTGAATGGCA 3180
TGGTGCATGT GATCACAGAC ATCCAGGTTG GATCCAGAAT CAACTATTCT TGTACTACAG 3240
```


cDNA sequence *continued*

```
GGCACCGACT CATTGGTCAC TCATCTGCTG AATGTATCCT CTCAGGCAAT ACTGCCCATT 3300
GGAGCACGAA GCCGCCAATT TGTCAACGAA TTCCTTGTGG GCTACCCCCA ACCATCGCCA 3360
ATGGAGATTT CATTAGCACC AACAGAGAGA ATTTTCACTA TGGATCAGTG GTGACCTACC 3420
GCTGCAATCT TGGAAGCAGA GGGAGAAAGG TGTTTGAGCT TGTGGGTGAG CCCTCCATAT 3480
ACTGCACCAG CAATGACGAT CAAGTGGGCA TCTGGAGCGG CCCCGCCCCT CAGTGCATTA 3540
TACCTAACAA ATGCACGCCT CCAAATGTGG AAAATGGAAT ATTGGTATCT GACAACAGAA 3600
GCTTATTTTC CTTAAATGAA GTTGTGGAGT TTAGGTGTCA GCCTGGCTTT GTCATGAAAG 3660
GACCCCGCCG TGTGAAGTGC CAGGCCCTGA ACAAATGGGA GCCAGAGTTA CCAAGCTGCT 3720
CCAGGGTGTG TCAGCCGCCT CCAGAAATCC TGCATGGTGA GCATACCCCA AGCCATCAGG 3780
ACAACTTTTC ACCTGGGCAG GAAGTGTTCT ACAGCTGTGA GCCTGGCTAT GACCTCAGAG 3840
GGGCTGCGTC TCTGCACTGC ACACCCCAGG GAGACTGGAG CCCTGAAGCC CCGAGATGTG 3900
CAGTGAAATC CTGTGATGAC TTCTTGGGTC AACTCCCTCA TGGCCGTGTG CTATTTCCAC 3960
TTAATCTCCA GCTTGGGGCA AAGGTGTCCT TTGTCTGTGA TGAAGGGTTT CGCTTAAAGG 4020
GCAGTTCCGT TAGTCATTGT GTCTTGGTTG GAATGAGAAG CCTTTGGAAT AACAGTGTTC 4080
CTGTGTGTGA ACATATCTTT TGTCCAAATC CTCCAGCTAT CCTTAATGGG AGACACACAG 4140
GAACTCCCTC TGGAGATATT CCCTATGGAA AAGAAATATC TTACACATGT GACCCCCACC 4200
CAGACAGAGG GATGACCTTC AACCTCATTG GGGAGAGCAC CATCCGCTGC ACAAGTGACC 4260
CTCATGGGAA TGGGGTTTGG AGCAGCCCTG CCCCTCGCTG TGAACTTTCT GTTCGTGCTG 4320
GTCACTGTAA AACCCCAGAG CAGTTTCCAT TTGCCAGTCC TACGATCCCA ATTAATGACT 4380
TTGAGTTTCC AGTCGGGACA TCTTTGAATT ATGAATGCCG TCCTGGGTAT TTTGGGAAAA 4440
TGTTCTCTAT CTCCTGCCTA GAAAACTTGC TCTGGTCAAG TGTTGAAGAC AACTGTAGAC 4500
GAAAATCATG TGGACCTCCA CCAGAACCCT TCAATGGAAT GGTGCATATA AACACAGATA 4560
CACAGTTTGG ATCAACAGTT AATTATTCTT GTAATGAAGG GTTTCGACTC ATTGGTTCCC 4620
CATCTACTAC TTGTCTCGTC TCAGGCAATA ATGTCACATG GGATAAGAAG GCACCTATTT 4680
GTGAGATCAT ATCTTGTGAG CCACCTCCAA CCATATCCAA TGGAGACTTC TACAGCAACA 4740
ATAGAACATC TTTTCACAAT GGAACGGTGG TAACTTACCA GTGCCACACT GGACCAGATG 4800
GAGAACAGCT GTTTGAGCTT GTGGGAGAAC GGTCAATATA TTGCACCAGC AAAGATGATC 4860
AAGTTGGTGT TTGGAGCAGC CCTCCCCCTC GGTGTATTTC TACTAATAAA TGCACAGCTC 4920
CAGAAGTTGA AAATGCAATT AGAGTACCAG GAAACAGGAG TTTCTTTTCC CTCACTGAGA 4980
TCATCAGATT TAGATGTCAG CCCGGGTTTG TCATGGTAGG GTCCCACACT GTGCAGTGCC 5040
AGACCAATGG CAGATGGGGG CCCAAGCTGC CACACTGCTC CAGGGTGTGT CAGCCGCCTC 5100
CAGAAATCCT GCATGGTGAG CATACCCTAA GCCATCAGGA CAACTTTTCA CCTGGGCAGG 5160
AAGTGTTCTA CAGCTGTGAG CCCAGCTATG ACCTCAGAGG GGCTGCGTCT CTGCACTGCA 5220
CGCCCCAGGG AGACTGGAGC CCTGAAGCCC CTAGATGTAC AGTGAAATCC TGTGATGACT 5280
TCCTGGGCCA ACTCCCTCAT GGCCGTGTGC TACTTCCACT TAATCTCCAG CTTGGGGCAA 5340
AGGTGTCCTT TGTTTGCGAT GAAGGGTTCC GATTAAAAGG CAGGTCTGCT AGTCATTGTG 5400
TCTTGGCTGG AATGAAAGCC CTTTGGAATA GCAGTGTTCC AGTGTGTGAA CAAATCTTTT 5460
GTCCAAATCC TCCAGCTATC CTTAATGGGA GACACACAGG AACTCCCTTT GGAGATATTC 5520
CCTATGGAAA AGAAATATCT TACGCATGCG ACACCCACCC AGACAGAGGG ATGACCTTCA 5580
ACCTCATTGG GGAGAGCTCC ATCCGCTGCA CAAGTGACCC TCAAGGGAAT GGGGTTTGGA 5640
GCAGCCCTGC CCCTCGCTGT GAACTTTCTG TTCCTGCTGC CTGCCCACAT CCACCCAAGA 5700
TCCAAAACGG GCATTACATT GGAGGACACG TATCTCTATA TCTTCCTGGG ATGACAATCA 5760
GCTACACTTG TGACCCCGGC TACCTGTTAG TGGGAAAGGG CTTCATTTTC TGTACAGACC 5820
AGGGAATCTG GAGCCAATTG GATCATTATT GCAAAGAAGT AAATTGTAGC TTCCCACTGT 5880
TTATGAATGG AATCTCGAAG GAGTTAGAAA TGAAAAAAGT ATATCACTAT GGAGATTATG 5940
TGACTTTGAA GTGTGAAGAT GGGTATACTC TGGAAGGCAG TCCCTGGAGC CAGTGCCAGG 6000
CGGATGACAG ATGGGACCCT CCTCTGGCCA AATGTACCTC TCGTGCACAT GATGCTCTCA 6060
TAGTTGGCAC TTTATCTGGT ACGATCTTCT TTATTTTACT CATCATTTTC CTCTCTTGGA 6120
TAATTCTAAA GCACAGAAAA GGCAATAATG CACATGAAAA CCCTAAAGAA GTGGCTATCC 6180
ATTTACATTC TCAAGGAGGC AGCAGCGTTC ATCCCCGAAC TCTGCAAACA AATGAAGAAA 6240
ATAGCAGGGT CCTTCCTTGA CAAAGTACTA TACAGCTGAA GAACATCTCG AATACAATTT 6300
TGGTGGGAAA GGAGCCAATT GATTTCAACA GAATCAGATC TGAGCTTCAT AAAGTCTTTG 6360
AAGTGACTTC ACAGAGACGC AGACATGTGC ACTTGAAGAT GCTGCCCCTT CCCTGGTACC 6420
TAGCAAAGCT CCTGCCTCTT TGTGTGCGTC ACTGTGAAAC CCCCACCCTT CTGCCTCGTG 6480
```

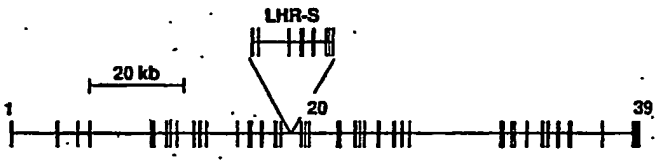
cDNA sequence *continued*

```
CTAAACGCAC ACAGTATCTA GTCAGGGGAA AAGACTGCAT TTAGGAGATA GAAAATAGTT 6540
TGGATTACTT AAAGGAATAA GGTGTTGCCT GGAATTTCTG GTTTGTAAGG TGGTCACTGT 6600
TCTTTTTTAA AATATTTGTA ATATGGAATG GGCTCAGTAA GAAGAGCTTG GAAAATGCAG 6660
AAAGTTATGA AAAATAAGTC ACTTATAATT ATGCTACCTA CTGATAACCA GTCCTAATAT 6720
TTTGATTCAT TTTCTGCCTA TCTTCTTTCA CATATGTGTT TTTTTACATA CGTACTTTTC 6780
CCCCCTTAGT TTGTTTCCTT TTATTTTATA GAGCAGAACC CTAGTCTTTT AAACAGTTTA 6840
GAGTGAAATA TATGCTATAT CAGTTTTTAC TTTCTCTAGG GAGAAAAATT AATTTACTAG 6900
AAAGGCATGA AATGATCATG GGAAGAGTGG TTAAGACTAC TGAAGAGAAA TATTTGGAAA 6960
ATAAGATTTC GATATCTTCT TTTTTTTTGA GATGGAGTCT GGCTCTGTCT CCCAGGCTGG 7020
AGTGCAGTGG CGTAATCTCG GCTCACTGCA ACGTCCGCCT CCTGGGTTGA CACCATTTTC 7080
CTGCCTCAGC CTCCTGAGTA GTTGGGACTA CCAGTAGATG GGACTACAGG CACCTGCCAA 7140
CACGCCCGGC TAATTTTTTT GTATTTTTAG TAGAGACGGG GTTTCACCAT GTTAGCCAGG 7200
ATGGTCTGGA TCTCCTGACC TCGTGATCCA CCCGCCTCGG CCTCCCAAAG TGCTGCGATT 7260
ACAGGCATGA GCCACCGCGC CTGGCCGCTT TCGATATTTT CTAAACTTTA ATTCAAAAGC 7320
ACTTTGTGCT GTGTTCTATA TAAAAAACAT AATAAAAATT GAAATGAAAG AATAATTGTT 7380
ATTATAAAAG TACTAGCTTA CTTTTGTATG GATTCAGAAT ATACTAAATT AACTTTTTAA 7440
AACACAACTT TTAAAAAATG TATCAAAAAT AATAAACGTG TTCTGATATT TTT
```

The first five nucleotides in each exon are underlined. There are two transcriptional start sites T1 and A30, the A is predominant by S1 nuclease analysis[60]. The two possible methionine initiation codons (ATG), the termination codon (TGA) and the known polyadenylation site (AATAAA) are indicated. In this figure, nucleotides 116–7061 are a compilation from references 3 and 4, determined from cDNA clones. Nucleotides 1–115 and 7062–7493 were determined from genomic clones[60].

Genomic structure[60]

The gene for the CR1*1 allotype of CR1 spans approximately 133 kb and is encoded by 39 exons as illustrated below.



The difference between the major allotypes is accounted for by deletion or duplication of a large segment of genomic DNA encoding an LHR-length of peptide sequence. The gene encoding the CR1*2 allotype is approximately 150–160 kb and is encoded by 47 exons, with the additional 8 exons inserted approximately in the location indicated. The gene encoding the CR1*3 allotype contains a deletion somewhere within the LHR-B to LHR-C regions, however the location has not been determined precisely[16].

Accession numbers (EMBL/GenBank)

		cDNA	Genomic
Human	CR1[3,4,60,60]	Y00816	L17390–L17430
Chimpanzee	CR1[61,66]	L24920–L24922	
Baboon	CR1[63]	L39791	
Mouse	CR1/CR2[62,63–68]	M61132	
		M36470	
		M29281	
		M35684	
		J04153	
		M33527	
		U17123–U17128	
		X98171	
Mouse	Crry[69]	M23529	
		M34164–M34173	
Rat	Crry[70]	L36532	
		D42115	

Deficiency

No humans totally lacking CR1 have been identified. The Knops, McCoy, Swain-Langley and York blood group antigens are located on CR1, and some individuals with these antibodies have very low levels of erythrocyte CR1[71]. Acquired low levels of erythrocyte CR1 are seen in patients with systemic lupus erythematosus[42–45]. These patients have abnormal clearance of immune complexes. Knockout mice have been prepared that lack CR1/CR2 and these animals exhibit profound defects in T cell and B cell function[21,72].

Polymorphic variants

The structural allotypes below are a consequence of large insertions or deletions in the CR1 gene, and may be detected by *M*, difference upon SDS-PAGE[5–6], northern blot analysis of mRNA or southern blot analysis of genomic DNA. The structural allotype may affect affinity of CR1 for C3b dimers[16]. The quantitative allotype, H or L, regulates CR1 expression level on erythrocytes. Erythrocytes from individuals homozygous for the H allotype bear 4–10-fold more cell surface CR1 than those from individuals homozygous for the L allotype[73].

Polymorphism frequencies[73–82]

Structural alleles	White population	African-Americans	Mexican-Americans	Chinese/Taiwanese
CR1*1	0.86–0.93	0.82–0.84	0.89	0.96
CR1*2	0.07–0.26	0.11–0.12	0.11	0.03
CR1*3	0–0.02	0.04–0.06	0	0.01
CR1*4	<0.01	<0.01	<0.01	0

Quantitative alleles	White population	African-Americans	Mexican-Americans	Chinese/ Taiwanese
H	0.75–0.78	0.85	0.80	0.71
L	0.25–0.22	0.15	0.20	0.28
Knops Phenotype	**White population**	**African-Americans**	**Mexican-Americans**	**Chinese/ Taiwanese**
Kn(a+)	0.98	0.99	Unknown	0.99
McC(a+)	0.98	0.94	Unknown	1.00
McC(b+)	0.01	0.51	Unknown	0.02
Sl(a+)	0.99	0.65	Unknown	0.97
Yk(a+)	0.98	0.92	Unknown	Unknown

References

1 Fearon, D.T. (1979) Proc. Natl Acad. Sci. USA 76, 5867–5871.
2 Fearon, D.T. (1980) J. Exp. Med. 152, 20–30.
3 Klickstein, L.B. et al. (1987) J. Exp. Med. 165, 1095–1112.
4 Klickstein, L.B. et al. (1988) J. Exp. Med. 168, 1699–1717.
5 Holers, V.M. et al. (1986) Complement 3, 63–78.
6 Dykman, T.R. et al. (1983) Proc. Natl Acad. Sci. USA 80, 1698–1702.
7 Wong, W.W. et al. (1983) J. Clin. Invest. 72, 685–693.
8 Dykman, T.R. et al. (1984) J. Exp. Med. 159, 691–703.
9 Dykman, T.R. et al. (1985) J. Immunol. 134, 1787–1789.
10 Wong, W.W. et al. (1985) J. Immunol. Methods 82, 303–313.
11 Lublin, D.M. et al. (1986) J. Biol. Chem. 261, 5736–5744.
12 Atkinson, J.M. and Jones, E.A. (1984) J. Clin. Invest. 74, 1649–1657.
13 Sim R.B. (1985) Biochem. J. 232, 883–889.
14 Lozier, J. et al. (1984) Proc. Natl Acad. Sci. USA 81, 3640–3644.
15 Weisman, H.F. et al. (1990) Science 249, 146–151.
16 Wong, W.W. and Farrell, S.A. (1991) 146, 656–662.
17 Klickstein, L.B. et al. (1997) Immunity 7, 345–355.
18 Kalli, K.R. et al. (1991) J. Immunol. 147, 590–594.
19 Nelson, R.A. (1953) Science 118, 733–737.
20 Bennaceraf, B. et al. (1959) J. Immunol. 82, 131–137.
21 Arend, W.P. and Mannik, M. (1971) J. Immunol. 107, 63–75.
22 Cornacoff, J.B. et al. (1983) J. Clin. Invest. 71, 236–247.
23 Schifferli, J.A. et al. (1988) J. Immunol. 140, 899–904.
24 Kimberly, R.P. et al. (1989) J. Clin. Invest. 84, 962–970.
25 Fearon, D.T. et al. (1981) J. Exp. Med. 153, 1615–1628.
26 Newman, S.L. et al. (1980) J. Immunol. 125, 2236–2244.
27 Wright, S.D. et al. (1983) J. Exp. Med. 158, 1338–1343.
28 Klaus, G.G.B. et al. (1980) Immunol. Rev. 53, 3–28.
29 Croix, D.A. et al. (1996) J. Exp. Med. 183, 1857–1864.
30 Carroll, M.C. (1998) Annu. Rev. Immunol. 16, 545–568.
31 Fang, Y. et al. (1998) J. Immunol. 160, 5273–5279.
32 Molina, H. et al. (1996) Proc. Natl Acad. Sci. USA 93, 3357–3361.
33 Iida, K. and Nussenzweig, V. (1981) J. Exp. Med. 153, 1138–1150.
34 Medof, M. and Nussenzweig, V. (1984) J. Exp. Med. 159, 1669–1685.
35 Kinoshita, T. et al. (1986) J. Exp. Med. 164, 1377–1388.

36 Fischer, E. et al. (1986) J. Immunol. 136, 1373–1377.
37 Wilson, J.G. et al. (1983) J. Immunol. 131, 684–689.
38 Tedder, T.F. et al. (1983) J. Immunol. 130, 1668–1673.
39 Yoon, S.H. and Fearon, D.T. (1985) J. Immunol. 134, 3332–3338.
40 Pascual, M. et al. (1993) J. Immunol. 151, 1702–1711.
41 Nickells, M. et al. (1998) Clin. Exp. Immunol. 112, 27–33.
42 Miyakawa, Y. et al. (1981) Lancet 2, 493–497.
43 Wilson, J.G. et al. (1982) N. Engl. J. Med. 307, 981–986.
44 Iida, K. et al. (1982) J. Exp. Med. 155, 1427–1438.
45 Walport, M.J. et al. (1985) Clin. Exp. Immunol. 59, 547–554.
46 Fearon, D.T. and Collins, L.A. (1983) J. Immunol. 130, 370–375.
47 Berger, M. et al. (1984) J. Clin. Invest. 74, 1566–1571.
48 O'Shea, J.J. et al. (1985) J. Immunol. 134, 2580–2587.
49 Berger, M. et al. (1991) Proc. Natl Acad. Sci. USA 88, 3019–3023.
50 Hourcade, D. et al. (1988) J. Exp. Med. 168, 1255–1270.
51 Krych, M. et al. (1991) Proc. Natl Acad. Sci. USA 88, 4353–4357.
52 Kalli, K.R. et al. (1991) J. Exp. Med. 174, 1451–1460.
53 Makrides, S.C. et al. (1992) J. Biol. Chem. 267, 24754–24761.
54 Krych, M. et al. (1994) J. Biol. Chem. 269, 13273–13278.
55 Reilly, B.D. et al. (1994) J. Biol. Chem. 269, 7696–7701.
56 Rey-Campos, J. et al. (1988) J. Exp. Med. 167, 664–669.
57 Carroll, M.C. et al. (1988) J. Exp. Med. 167, 1271–1280.
58 Kingsmore, S.F. et al. (1989) J. Exp. Med. 169, 1479–1484.
59 Kurtz, C.B. et al. (1989) J. Immunol. 143, 2058–2067.
60 Vik, D.P. and Wong, W.W. (1993) J. Immunol. 151, 6214–6224.
61 Nickells, M.W. et al. (1995) J. Immunol. 154, 2829–2837.
62 Molina, H. et al. (1990) J. Immunol. 145, 2974–2983.
63 Kinoshita, T. et al. (1988) J. Immunol. 140, 3066–3072.
64 Birmingham, D.J. et al. (1994) J. Immunol. 153, 691–700.
65 Clemenza, L. et al. (1997) Mol. Immunol. 34, 297–304.
66 Molina, H. et al. (1990) J. Immunol. 145, 2974–2983.
67 Kurtz, C.B. et al. (1990) J. Immunol. 144, 3581–3591.
68 Fingeroth, J.D. (1990) J. Immunol. 144, 3458–3467.
69 Paul, M.S. et al. (1989) J. Immunol. 142, 582–589.
70 Quigg, R.J. et al. (1995) Immunogenet. 42, 362–367.
71 Moulds, J.M. et al. (1991) J. Exp. Med. 173, 1159–1163.
72 Ahearn, J.M. et al. (1996) Immunity 4, 251–262.
73 Wilson, J.G. et al. (1986) J. Exp. Med. 164, 50–59.
74 Van Dyne, S. et al. (1987) Clin. Exp. Immunol. 68, 570–579.
75 Wong, W.W. et al. (1986) J. Exp. Med. 164, 1531–1546.
76 ATCC Product #57330 (*E. coli*) or #57331 (plasmid DNA).
77 Wong, W.W. et al. (1985) Proc. Natl Acad. Sci. USA 82, 7711–7715.
78 Moldenhauer, F. (1987) Arthritis Rheum. 30, 961–966.
79 Cohen, J.H.M. et al. (1989) Arthritis Rheum. 32, 393–398.
80 Tebib, J.G. et al (1989) Arthritis Rheum. 32, 1465–1469.
81 Moulds, J.M. et al. (1996) Clin. Exp. Immunol. 105, 302–305.
82 Moulds, J.M. et al. (1998) Exp. Clin. Immunogenet. 15, 291–294.

Decay-accelerating factor

CD55, DAF

L. Kuttner-Kondo, W.G. Brodbeck, and M.E. Medof, Institute of Pathology, Case Western Reserve University, Cleveland, OH, USA

Physicochemical properties

Decay-accelerating factor is synthesized as a single-chain prepropeptide of 381 amino acids including a 34 amino acid leader sequence, and a C-terminal signal sequence of 28 amino acids which is cleaved and replaced with a glycosylphosphatidylinositol (GPI) anchor[1,2].

M_r (K)[3,4]	70 (non-reduced)
	75 (reduced)
N-linked glycosylation sites	1 (95)[1,2,5]

Structure

DAF is a glycoprotein composed of four CCP domains, a heavily glycosylated STP (membrane proximal) domain rich in serines, threonines and prolines, and a GPI anchor[1,2,6].

Function

DAF functions intrinsically in cell membranes to protect host cells from autologous complement attack[4]. It accelerates the decay of the classical and alternative C3 and C5 convertases[3,4,7,8]. The purified protein, when added to cells, incorporates into their membranes and is functional[4].

Tissue distribution

DAF is present on all blood elements[9,10], and most other cell types. It is expressed in high levels on cells that line extravascular compartments[11]. Soluble variants of DAF are found in body fluids[11,12].

Regulation of expression[13–15]

Varies from 2500 molecules/cell on erythrocytes to >100 000 molecules/cell on endothelial and epithelial cells. Induced by phorbol ester (PMA). Promoter has cAMP response element.

Protein sequence[1,2]

```
MTVARPSVPA ALPLLGELPR LLLLVLLCLP AVWGDCGLPP DVPNAQPALE   50
GRTSFPEDTV ITYKCEESFV KIPGEKDSVI CLKGSQWSDI EEFCNRSCEV  100
PTRLNSASLK QPYITQNYFP VGTVVEYECR PGYRREPSLS PKLTCLQNLK  150
WSTAVEFCKK KSCPNPGEIR NGQIDVPGGI LFGATISFSC NTGYKLFGST  200
SSFCLISGSS VQWSDPLPEC REIYCPAPPQ IDNGIIQGER DHYGYRQSVT  250
YACNKGFTMI GEHSIYCTVN NDEGEWSGPP PECRGKSLTS KVPPTVQKPT  300
TVNVPTTEVS PTSQKTTTKT TTPNAQATRS TPVSRTTKHF HETTPNKGSG  350
TTSGTTRLLS GHTCFTLTGL LGTLVTMGLL T
```

The leader sequence is underlined, *N*-linked glycosylation site is indicated (N), and the cleavage site for GPI anchor attachment is double underlined. Amino acid differences between the two publications: 80 I/T, 85 S/M.

Protein modules

1–34	Leader sequence[1,2]	exon 1[16]
35–95	CCP	exon 2
97–159	CCP	exon 3
162–221	CCP	exon 4/5
224–284	CCP	exon 6
287–356	STP	exon 7–9

360

Chromosomal location

Human[17]: 1q32.
Mouse[18]: chromosome 1.

cDNA sequence[1,2,13]

```
ACTGCAACTC GCTCCGGCCG CTGGGCGTAG CTGCGACTCG GCGGAGTCCC GGCGGCGCGT   60
CCTTGTTCTA ACCCGGCGCG CCATGACCGT CGCGCGGCCG AGCGTGCCCG CGGCGCTGCC  120
CCTCCTCGGG GAGCTGCCCC GGCTGCTGCT GCTGGTGCTG TTGTGCCTGC CGGCCGTGTG  180
GGGTGACTGT GGCCTTCCCC CAGATGTACC TAATGCCCAG CCAGCTTTGG AAGGCCGTAC  240
AAGTTTCCC GAGGATACTG TAATAACGTA CAAATGTGAA GAAAGCTTTG TGAAAATTCC  300
TGGCGAGAAG GACTCAGTGA TCTGCCTTAA GGGCAGTCAA TGGTCAGATA TTGAAGAGTT  360
CTGCAATCGT AGCTGCGAGG TGCCAACAAG GCTAAATTCT GCATCCCTCA AACAGCCTTA  420
TATCACTCAG AATTATTTTC CAGTCGGTAC TGTTGTGGAA TATGAGTGCC GTCCAGGTTA  480
CAGAAGAGAA CCTTCTCTAT CACCAAAACT AACTTGCCTT CAGAATTTAA AATGGTCCAC  540
AGCAGTCGAA TTTTGTAAAA AGAAATCATG CCCTAATCCG GGAGAAATAC GAAATGGTCA  600
GATTGATGTA CCAGGTGGCA TATTATTTGG TGCAACCATC TCCTTCTCAT GTAACACAGG  660
GTACAAATTA TTTGGCTCGA CTTCTAGTTT TTGTCTTATT TCAGGCAGCT CTGTCCAGTG  720
GAGTGACCCG TTGCCAGAGT GCAGAGAAAT TTATTGTCCA GCACCACCAC AAATTGACAA  780
TGGAATAATT CAAGGGGAAC GTGACCATTA TGGATATAGA CAGTCTGTAA CGTATGCATG  840
TAATAAAGGA TTCACCATGA TTGGAGAGCA CTCTATTTAT TGTACTGTGA ATAATGATGA  900
AGGAGAGTGG AGTGGCCCAC CACCTGAATG CAGAGGAAAA TCTCTAACTT CCAAGGTCCC  960
ACCAACAGTT CAGAAACCTA CCACAGTAAA TGTTCCAACT ACAGAAGTCT CACCAACTTC 1020
TCAGAAAACC ACCACAAAAA CCACCACACC AAATGCTCAA GCAACACGGA GTACACCTGT 1080
TTCCAGGACA ACCAAGCATT TTCATGAAAC AACCCCAAAT AAAGGAAGTG GAACCACTTC 1140
AGGTACTACC CGTCTTCTAT CTGGGCACAC GTGTTTCACG TTGACAGGTT TGCTTGGGAC 1200
GCTAGTAACC ATGGGCTTGC TGACTTAGCC AAAGAAGAGT TAAGAAGAAA ATACACACAA 1260
GTATACAGAC TGTTCCTAGT TTCTTAGACT TATCTGCATA TTGGATAAAA TAAATGCAAT 1320
TGTGCTCTTC ATTTAGGATG CTTTCATTGT CTTTAAGATG TGTTAGGAAT GTCAACAGAG 1380
CAAGGAGAAA AAAGGCAGTC CTGGAATCAC ATTCTTAGCA CACCTACACC TCTTGAAAAT 1440
AGAACAACTT GCAGAATTGA GAGTGATTCC TTTCCTAAAA GTGTAAGAAA GCATAGAGAT 1500
TTGTTCGTAT TTAGAATGGG ATCACGAGGA AAAGAGAAGG AAAGTGATTT TTTTCCACAA 1560
GATCTGTAAT GTTATTTCCA CTTATAAAGG AAATAAAAAA TGAAAAACAT TATTTGGATA 1620
TCAAAAGCAA ATAAAAACCC AATTCAGTCT CTTCTAAGCA AAATTGCTAA AGAGAGATGA 1680
ACCACATTAT AAAGTAATCT TTGGCTGTAA GGCATTTTCA TCTTTCCTTC GGGTTGGCAA 1740
AATATTTTAA AGGTAAAACA TGCTGGTGAA CCAGGGGTGT TGATGGTGAT AAGGGAGGAA 1800
TATAGAATGA AAGACTGAAT CTTCCTTTGT TGCACAAATA GAGTTTGGAA AAAGCCTGTG 1860
AAAGGTGTCT TCTTTGACTT AATGTCTTTA AAAGTATCCA GAGATACTAC AATATTAACA 1920
TAAGAAAAGA TTATATATTA TTTCTGAATC GAGATGTCCA TAGTCAAATT TGTAAATCTT 1980
ATTCTTTTGT AATATTTATT TATATTTATT TATGACAGTG AACATTCTGA TTTTACATGT 2040
AAAACAAGAA AAGTTGAAGA AGATATGTGA AGAAAAATGT ATTTTTCCTA AATAGAATTA 2100
AATGATCCCA TTTTTTGGT
```

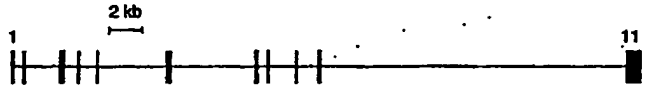
Position 1 is the transcriptional start site. The first five nucleotides in each exon are underlined. Exon 10 (not illustrated in the cDNA sequence but depicted below), an *Alu* family sequence, has not been reported in DAF mRNA and is not used in surface DAF protein[1]. The initiation codon (ATG), the termination codon (TAG), and the four polyadenylation signals (AATAAA) are indicated. Nucleotide differences between the published sequences: 321 T/C, 336 G/T, 337 T/G.

Exon 10[1]

1164 GTTCTCGTCC TGTCACCCAG GCTGGTATGC GGTGGTGTGA TCGTAGCTCA CTGCAGTCTC
GAACTCCTGG GTTCAAGCGA TCCTTCCACT TCAGCCTCCC AAGTAGCTGG TACTACAG

Genomic structure[26]

The gene spans ~40 kb and is encoded by 11 exons. The introns vary in size from 0.5 to19.8 kb (last intron). Exon 10 has not been reported in DAF mRNA.



Accession numbers (EMBL/GenBank)

Human	M31516 (Caras)	M64356 (promoter)
	M15799 (Medof)	
	M64653	
	S72858	
Mouse[14,19]	L41365–6	
	D63679	
Orang-utan[20]	S67775	
Guinea-pig[21]	D49416–D49421	
Rat[22]	AF039583–4	

Deficiency

Paroxysmal nocturnal haemoglobinuria results from the absence of DAF (as well as CD59 and all other GPI-anchored proteins) on peripheral blood elements[10,23,24]. The failure to express these proteins is due to a defect in the first step of GPI assembly (GlcNAc-PI synthesis[25–27]) in a bone marrow stem cell eventuating from a mutation of the *PIG-A* gene[28]. Deficient expression of DAF gives rise to heightened uptake of C3b[29]. The Cromer blood group antigen system resides on the DAF molecule[30]; the Inab phenotype represents the absence of DAF[31].

Polymorphic variants[32–36]

G237T; R52L
G237C; R52P
T327G; L82R
C678T; S199L
G761C; A227P
T321A; I80N
C831T; T250M

References

1 Caras, I.W. et al. (1987) Nature 325, 545–549.
2 Medof, M.E. et al. (1987) Proc. Natl Acad. Sci. USA 84, 2007–2011.
3 Nicholson-Weller, A. et al. (1982) J. Immunol. 129, 184–189.

4 Medof M.E. et al. (1984) J. Exp. Med. 160, 1558–1578.
5 Lublin, D.M. et al. (1986) J. Immunol. 137, 1629–1635.
6 Medof, M.E. et al. (1986) Biochemistry 25, 6740–6747.
7 Pangburn, M.K. (1986) J. Immunol. 136, 2216–2221.
8 Fujita, T. et al. (1987) J. Exp. Med. 166, 1221–1228.
9 Kinoshita, T. et al. (1985) J. Exp. Med. 162, 75–92.
10 Nicholson-Weller, A. et al. (1985) Blood 65, 1237–1244.
11 Medof, M.E. et al. (1987) J. Exp. Med. 165, 848–864.
12 Lass, J.H. et al. (1990) Invest. Ophthalmol. Vis. Sci. 31, 1136–1148.
13 Ewulonu, U.K. et al. (1991) Proc. Natl Acad. Sci. USA 88, 4675–4679.
14 Thomas, D.J. and Lublin, D.M. (1993) J. Immunol. 150, 151–160.
15 Bryant, R.W. et al. (1990) J. Immunol. 144, 593–598.
16 Post, T.W. et al. (1990) J. Immunol. 144, 740–744.
17 Lublin, D.M. et al. (1987) J. Exp. Med. 165, 1731–1736.
18 Spicer, A.P. et al. (1995) J. Immunol. 155, 3079–3091.
19 Fukuoka, Y. et al. (1996) Int. Immunol. 8, 379–385.
20 Nickells, M.W. et al. (1994) J. Immunol. 152, 676–685.
21 Nonaka, M. et al. (1995) J. Immunol. 155, 3037–3048.
22 Hinchliffe, S.J. et al. (1998). J. Immunol. 161, 5695–5703
23 Pangburn, M.K. et al. (1983) Proc. Natl Acad. Sci. USA 80, 5430–5434.
24 Nicholson-Weller, A. et al. (1983) Proc. Natl Acad. Sci. USA 80, 5066–5070.
25 Armstrong, C. et al. (1992) J. Biol. Chem. 267, 25347–25351.
26 Takahashi, M. et al. (1993) J. Exp. Med. 177, 517–521.
27 Hidaka, M. et al. (1993) Biochim. Biophys. Acta 191, 571–579.
28 Takeda, J. et al. (1993) Cell 73, 703–711.
29 Medof, M.E. et al. (1985) Proc. Natl Acad. Sci. USA 82, 2980–2984.
30 Telen, M.J. et al. (1988) J. Exp. Med. 167, 1993–1998.
31 Parsons, S.F. et al. (1988) Proceedings of the 20th Congress of the International Society of Blood Transfusion, London, UK, p. 116 (abstr.).
32 Stafford, H.A. et al. (1988) Proc. Natl Acad. Sci. USA 85, 880–884.
33 Telen, M.J. and Green, A.M. (1989) Blood 74, 437–441.
34 Lublin, D.M. et al. (1991) J. Clin. Invest. 87, 1945–1952.
35 Telen, M.J. et al. (1994) Blood 84, 3205–3211.
36 Lublin, D.M. et al. (1997) Transfusion 37, 102S (abstr.).

Membrane cofactor protein

M. Kathryn Liszewski and John P. Atkinson, Department of Medicine, Washington University School of Medicine, St Louis, MO, USA

Other names

MCP, CD46, gp45-70, measles virus receptor.

Physicochemical properties

MCP is a type 1 membrane glycoprotein expressed as four common isoforms (C1, C2, BC1 and BC2) that arise by alternative splicing[1]. Each possesses a 34 amino acid signal peptide followed by 328–350 amino acids (C1, 328 aa; C2, 335 aa; BC1, 343 aa; BC2, 350 aa)[2–4].

pI		3.9–5.8
Higher molecular weight species possess more O-linked sugars, which correlate with their more acidic pI.		
M_r (K)	predicted	~39
	observed	51–58 (C isoforms)
		59–68 (BC isoforms)
N-linked glycosylation sites[2,5] (all occupied)		3 (83, 114, 273)

Structure

The N-terminal portion of all isoforms consists of four contiguous CCP repeats. Following the four CCP is an alternatively spliced segment enriched in serines, threonines and prolines (STP domain) that is O-glycosylated at positions 286–314 (BC isoforms) and 286–299 (C isoforms)[2,5]. The gene contains three STP exons termed A, B and C. Two of the regularly expressed isoforms consist of B+C exons (29 amino acids) and two consist of the C (14 amino acids) exon only. STP exon A is rarely used. Flanking the STP domain and common to all of the isoforms is a sequence of 12 amino acids of unknown function, which is followed by the transmembrane domain, and intracytoplasmic anchor. Alternative splicing produces two distinct cytoplasmic tails of 16 or 23 amino acids termed Cyt-1 or Cyt-2.

Designations for the four commonly occurring isoforms are MCP-BC1, MCP-BC2, MCP-C1 and MCP-C2 (specifying the alternatively spliced STP and tail regions). Other rarer isoforms of unclear significance have been described. The upper band on SDS-PAGE consists of BC1 and BC2 while the lower band consists of C1 and C2. Population studies indicate that the upper band predominant pattern is present in 65% of the population, an approximately equal distribution of upper and lower bands occurs in 29%, and 6% express the lower band (C isoforms) predominantly[6].

Function

MCP is a ubiquitously expressed complement regulatory protein. It is a cofactor for the factor I-mediated cleavage of C3b and C4b that deposits on self-tissue[7]. This regulation is only performed intrinsically in that MCP

protects the cell on which it is anchored, not neighbouring cells[8]. MCP is expressed on placental trophoblast and on the inner acrosomal membrane of human spermatozoa[9,10]. Its role in these locations is likely to protect against complement activation, but other possibilities have been suggested[11].

Crosslinking of MCP downregulates IL-12 production, a finding of potential significance for the immunosuppressive sequelae of measles virus infection[12].

MCP is a receptor for several pathogens including measles virus[12–15], group A *Streptococcus pyogenes*[16], and pathogenic *Neisseria*[17]. Therapeutic uses of MCP include production of transgenic animals expressing MCP in order to prevent the hyperacute graft rejection that accompanies xenotransplantation[18] and a recombinantly produced soluble form in which MCP is linked with a second complement regulatory protein (decay-accelerating factor) for therapeutic use as an inhibitor of complement activation[19].

Tissue distribution

Most cells express each of the four isoforms of MCP. Human erythrocytes lack MCP. Tissue-specific isoform expression has been found in kidney and salivary gland (BC isoforms) as well as brain and fetal heart (C isoforms)[20,21].

Regulation of expression

MCP levels are increased in certain haematologic malignancies, on most solid tumour cell lines, and following SV40 transformation (reviewed in ref. 22). MCP expression is upregulated in glomerular capillary walls and in mesangial regions of diseased kidney tissues and in astrocytes following cytomegalovirus infection. IFNγ and phorbal ester (PMA) enhanced expression in an oligodendrocyte cell line[23].

Protein sequence (BC2)[2]

MEPPGRRECP FPSWRFPGLL LAAMVLLLVS FSDACEEPPT FEAMELIGKP 50
KPYYEIGERV DYKCKKGYFY IPPLATHTIC DRNHTWLPVS DDACYRETCP 100
YIRDPLNGQA VPANGTYEFG YQMHFICNEG YYLIGEEILY CELKGSVAIW 150
SGKPPICEKV LCTPPPKIKN GKHTFSEVEV FEYLDAVTYS CDPAPGPDPF 200
SLIGESTIYC GDNSVWSRAA PECKVVKCRF PVVENGKQIS GFGKKFYYKA 250
TVMFECDKGF YLDGSDTIVC DSNSTWDPPV PKCLKVSTSS TTKSPASSAS 300
GPRPTYKPPV SNYPGYPKPE EGILDSLDVW VIAVIVIAIV VGVAVICVVP 350
YRYLQRRKKK GKADGGAEYA TYQTKSTTPA EQRG

The leader sequence is underlined, N-linked glycosylation sites (all occupied) are indicated (N) and segments alternatively spliced are double underlined (see Structure).

Protein modules

1–34	Leader peptide	exon 1
35–95	CCP	exon 2
96–158	CCP	exon 3/4
159–224	CCP	exon 5
225–285	CCP	exon 6
286–314	STP	exon 7–9
	B domain: VSTSSTTKSPASSAS	exon 8
	C domain: GPRPTYKPPVSNYP	exon 9
315–327	Undefined segment	exon 10
328–351	Transmembrane domain	exon 11/12
352–361	Intracytoplasmic anchor	exon 12
362–377	Cytoplasmic tail one: TYLTDETHREVKFTSL	exon 13
362–384	Cytoplasmic tail two: KADGGAEYATYQTKSTTPAEQRG	exon 14

Chromosomal location[2,24,25]

Human: 1q3.2.

It is located along with four other closely related genes on a 900 kb fragment within the RCA locus at 1q3.2. An MCP-like genomic element includes sequences 93% homologous at the nucleotide level with the MCP 5′ terminus (i.e. signal peptide, and CCP1–3)[25]. Located within 60 kb of MCP[25], it is unknown if this partial duplication produces a protein.

cDNA sequence (BC2)[2]

```
TCTGCTTTCC TCCGGAGAAA TAACAGCGTC TTCCGCGCCG CGCATGGAGC CTCCCGGCCG    60
CCGCGAGTGT CCCTTTCCTT CCTGGCGCTT TCCTGGGTTG CTTCTGGCGG CCATGGTGTT   120
GCTGCTGTAC TCCTTCTCCG ATGCCTGTGA GGAGCCACCA ACATTTGAAG CTATGGAGCT   180
CATTGGTAAA CCAAAACCCT ACTATGAGAT TGGTGAACGA GTAGATTATA AGTGTAAAAA   240
AGGATACTTC TATATACCTC CTCTTGCCAC CCATACTATT TGTGATCGGA ATCATACATG   300
GCTACCTGTC TCAGATGACG CCTGTTATAG AGAAACATGT CCATATATAC GGGATCCTTT   360
AAATGGCCAA GCAGTCCCTG CAAATGGGAC TTACGAGTTT GGTTATCAGA TGCACTTTAT   420
TTGTAATGAG GGTTATTACT TAATTGGTGA AGAAATTCTA TATTGTGAAC TTAAAGGATC   480
AGTAGCAATT TGGAGCGGTA AGCCCCCAAT ATGTGAAAAG GTTTTGTGTA CACCACCTCC   540
AAAAATAAAA AATGGAAAAC ACACCTTTAG TGAAGTAGAA GTATTTGAGT ATCTTGATGC   600
AGTAACTTAT AGTTGTGATC CTGCACCTGG ACCAGATCCA TTTTCACTTA TTGGAGAGAG   660
CACGATTTAT TGTGGTGACA ATTCAGTGTG GAGTCGTGCT GCTCCAGAGT GTAAAGTGGT   720
CAAATGTCGA TTTCCAGTAG TCGAAAATGG AAAACAGATA TCAGGATTTG GAAAAAAATT   780
TTACTACAAA GCAACAGTTA TGTTTGAATG CGATAAGGGT TTTTACCTCG ATGGCAGCGA   840
CACAATTGTC TGTGACAGTA ACAGTACTTG GGATCCCCCA GTTCCAAAGT GTCTTAAAGT   900
GTCGACTTCT TCCACTACAA AATCTCCAGC GTCCAGTGCC TCAGGTCCTA GGCCTACTTA   960
CAAGCCTCCA GTCTCAAATT ATCCAGGATA TCCTAAACCT GAGGAAGGAA TACTTGACAG  1020
TTTGGATGTT TGGGTCATTG CTGTGATTGT TATTGCCATA GTTGTTGGAG TTGCAGTAAT  1080
TTGTGTTGTC CCGTACAGAT ATCTTCAAAG GAGGAAGAAG AAAGGGAAAG CAGATGGTGG  1140
AGCTGAATAT GCCACTTACC AGACTAAATC AACCACTCCA GCAGAGCAGA GAGGCTGAAT  1200
AGATTCCACA ACCTGGTTTG CCAGTTCATC TTTTGACTCT ATTAAAATCT TCAATAGTTG  1260
TTATTCTGTA GTTTCACTCT CATGAGTGCA ACTGTGGCTT AGCTAATATT GCAATGTGGC  1320
```

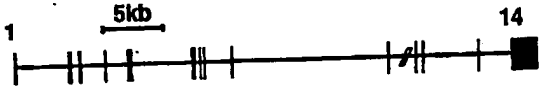

cDNA sequence (BC2) *continued*

```
TTGAATGTAG GTAGCATCCT TTGATGCTTC TTTGAAACTT GTATGAATTT GGGTATGAAC  1380
AGATTGCCTG CTTTCCCTTA AATAACACTT AGATTTATTG GACCAGTCAG CACAGCATGC  1440
CTGGTTGTAT TAAAGCAGGG ATATGCTGTA TTTTATAAAA TTGGCAAAAT TAGAGAAATA  1500
TAGTTCACAA TGAAATTATA TTTTCTTTGT
```

The first five nucleotides in each exon are underlined to indicate the intron–exon boundaries. The methionine initiation codon (ATG), the termination codon (TGA) and the probable polyadenylation signals (AATATA or AATGAA) are indicated.

Genomic structure

The gene spans a minimum of 43 kb and is encoded by 14 exons[3]. There are two sites for alternative splicing: exons 7, 8 and 9 encode the STP domains commonly expressed as isoforms with B+C (8 + 9) or C (9) alone; exons 13 and 14 encode the cytoplasmic tails, CYT-1 and CYT-2. Since exon 13 contains an in-frame stop codon, its expression as CYT-1 converts exon 14 into the 3′ untranslated region of MCP.



Accession numbers

Human[2,3]	MCP-BC2	Y00651
	MCP-BC1	X59405
	MCP-C1	X59406
	MCP-C2	X59407
	MCP-ABC2	X59409
	MCP-ABC1	X59410
Owl monkey[25]		U87914
Baboon[26]		U87915
Goeldii marmoset[26]		U87916
Common marmoset[26]		U87917
Tamarin[26]		U87918
Squirrel monkey[26]		U87919
African green monkey[26]		U87920
Cynomologous monkey[26]		U87921
Rhesus monkey[26]		U87922
White-faced saki[26]		U87923
Guinea-pig[27]		D84130–3
Pig[24,29]		D70897
Mouse[30]		AB001566

Deficiency

None known.

Polymorphic variants

A *Hind*III RFLP has been found that correlates with the phenotypic polymorphism of MCP[24]. This size polymorphism results from variable splicing of exon 8[3]. *Pvu*II and *Bgl*II RFLPs have also been described[21,22].

References

1 Liszewski, M.K. et al. (1996) Adv. Immunol. 61, 201–283.
2 Lublin, D.M. et al. (1988) J. Exp. Med. 168, 181–194.
3 Post, T.W. et al. (1991) J. Exp. Med. 174, 93–102.
4 Russell, S.M. et al. (1992) Eur. J. Immunol. 22, 1513–1518.
5 Ballard, L.L. et al. (1988) J. Immunol. 141, 3923–3929.
6 Ballard, L. et al. (1987) J. Immunol. 138, 3850–3855.
7 Seya, T. and Atkinson, J.P. (1989) Biochem. J. 264, 581–588.
8 Oglesby, T.J. et al. (1992) J. Exp. Med. 175, 1547–1551.
9 Cervoni, F. et al. (1992) J. Immunol. 148, 1431–1437.
10 Anderson, D.J. et al. (1989) Biol. Reprod. 41, 285–293.
11 Anderson, D.J. et al. (1993) Proc. Natl Acad. Sci. USA 90, 10051–10055.
12 Karp, C.L. et al. (1996) Science 273, 228–231.
13 Naniche, D. et al. (1993) J. Virol. 67, 6025–6032.
14 Dorig, R.E. et al. (1993) Cell 75, 295–305.
15 Manchester, M. et al. (1994) Proc. Natl Acad. Sci. USA 91, 2161–2165.
16 Okada, N. et al. (1995) Proc. Natl Acad. Sci. USA 92, 2489–2493.
17 Kallstrom, H. et al. (1997) Mol. Microbiol. 25, 639–647.
18 Cozzi, E. and White, D.J.G. (1995) Nature Med. 1, 964–966.
19 Higgins, P.J. et al. (1997) J. Immunol. 158, 2872–2881.
20 Johnstone, R.W. et al. (1993) Mol. Immunol. 30, 1231–1241.
21 Gorelick, A. et al. (1995) Lupus 4, 293–296.
22 Liszewski, M.K. et al. (1998) In (Rother, K., Till, G.O. and Hansch, G.M. eds). 2nd ed. Berlin, Springer-Verlag, pp. 146–162.
23 Gasque, P. and Morgan, B.P. (1996) Immunology 89, 338–347.
24 Bora, N.S. et al. (1989) J. Exp. Med. 169, 597–602.
25 Hourcade, D. et al. (1992) Genomics 12, 289–300.
26 Hsu, E.C. et al. (1997) J. Virol. 71, 6144–6154.
27 Hosokawa, M. et al. (1996) J. Immunol. 157, 4946–4952.
28 van den Berg, C.W. et al. (1997) J. Immunol. 158, 1703–1709.
29 Toyomura, K. et al. (1997) Int. Immunol. 9, 869–876.
30 Tsujimura, A. et al. (1998) Biochem. J. 330, 163–168.
31 Bora, N.S. et al. (1991) J. Immunol. 146, 2821–2825.
32 Wilton, A.N. et al. (1992) Immunogenetics 36, 79–85.

Schematic Drawing of Hybrid and Chimeric Proteins

3I3(A):

OOOOSSPNKOOOOOOOOOOOOSSPAPRCGI-HHHHHH

DAF	CR1	6XHis Tag
CCP1,2,3,4	CCP4,5,6,7,8,9,10,11,12,13,14	

DAFBB(2):

OOOOSSPNKOOOOOØOOOOOOOØOOOOOSSPAPRCGI-HHHHHH

DAF	CR1	6XHis Tag
CCP1,2,3,4	CCP4,5,6,7,8,9/9,10,11,12,13,14,15,16/9,10,11,12,13,14	

Mini (I-1) (Dimer):

OOOOSSPNKOOOOCGILVE≡□□

‖

OOOOSSPNKOOOOCGILVE≡□□

DAF	CR1	IgG4
CCP1,2,3,4	CCP4,5,6,7	Hinge-CH2-CH3

Note: Valine (V) is the last amino acid of IgG4's CH1 domain.

Micki (5A):

OOOOSSPNKOOOOCGILGHOOOOVS-HHHHHH

DAF	CR1	MCP	6XHis Tag
CCP1,2,3,4	CCP4,5,6,7	CCP1,2,3,4 + 1st 2 AA OF STP REGION	

Note: Valine-Serine (VS) are the 1st 2 amino acids (AA) of MCP's STP region

KEY:

O = CCP; Ø = Split CCP; ≡□□ = IgG Hinge and Constant Heavy Regions 2 and 3

‖ = Disulfide Bridges forming a dimer

DAF-CR1 HYBRID and CHIMERIC DESIGNS

DAF-CR1 HYBRIDS:

3I3: (DAF1 module EcoRI to XhoI + CR828XN3 XhoI to NsiI + CR1300NBF NsiI to BglII cloned into pSG5 expression vector)

```
                pSG5---EcoRI-BglII-DAF(Signal, CCP1-4)-SSPNK-CR1(CCP4-14[SSPAPRCGI])-6xHis-Stop-BglII---pSG5
                                                ↓               ↓                 ↓            1055
4076bp          pSG5:  1043                                     ↓           2953 EcoRI (gaattc)
                                           XhoI(ctcgag)
                                                       1603 EcoRI (CCP7-8 Linker)
                                                       1666 NsiI (CCP8)
                                                       1860 BamHI (CCP9)
                                                       1978 EcoRI (CCP9-10 Linker)
                                                       2442 XmaI (CCP12)
                                                       2442 SmaI (CCP12)
```

DAFBB: (3I3 cut with BamHI and BamHI 1861-3210 fragment from CR1/AprM8 introduced)

```
pSG5---EcoRI-BglII-DAF(Signal, CCP1-4)-SSPNK-CR1(CCP4-9[CCP9-16]CCP9-14[SSPAPRCGI]-6xHis-Stop-BglII---pSG5
                                                ↓  ↓     ↓  ↓
                                       NsiI: 1666  ↓  3016  ↓
                                       BamHI:     1860     3210
```

DAF-CR1-MCP HYBRID:

Micki: EcoRI-BglII-DAF(Signal + CCP1,2,3,4)-SSPNK-CR1(CCP4,5,6,7)–EcoRI–MCP(CCP1,2,3,4)-VS-6xHis-Stop- Bgl II into pSG5

```
                                                          1603
                                                  tgt gga att ctg ggt cac
                                                   C   G   I   L   G   H
                                                          CR1
```

*VS = 1st two amino acids of MCP STP region which follows MCP's CCP4

44

DAF-CR1 HYBRID and CHIMERIC DESIGNS

DAF-CR1-IgG Chimera (IgG4 from pHC-huCg4 [gift of Gary McLean, 2222 Health Sciences Mall, Vancouver, B.C., Canada]):

Mini: EcoRI-BglII-DAF (Signal + CCP1,2,3,4)-SSPNK-CR1(CCP4,5,6,7) – EcoRI – IgG4 (Hinge-CH2-CH3)-Stop-Bgl II into pSG5

1603

tgt gga att ctg gtt gag

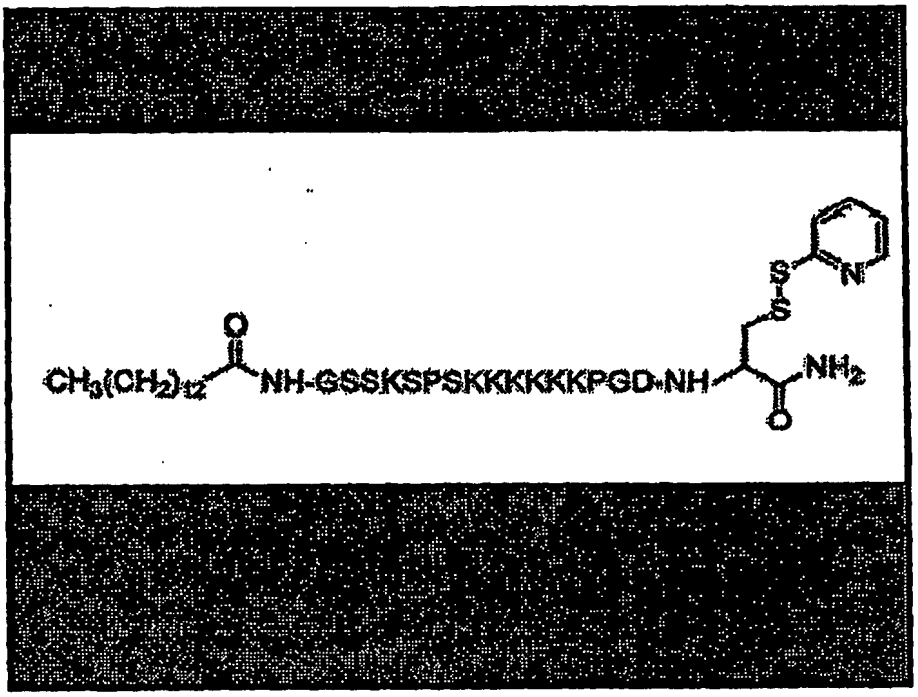
C G I L V E

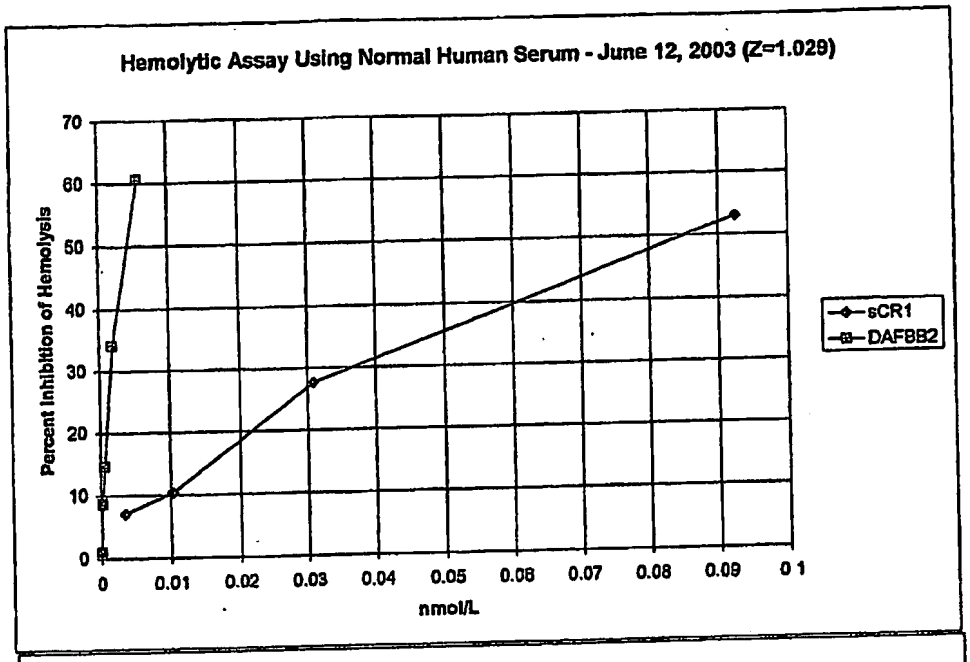
CR1 - IgG4 Linker

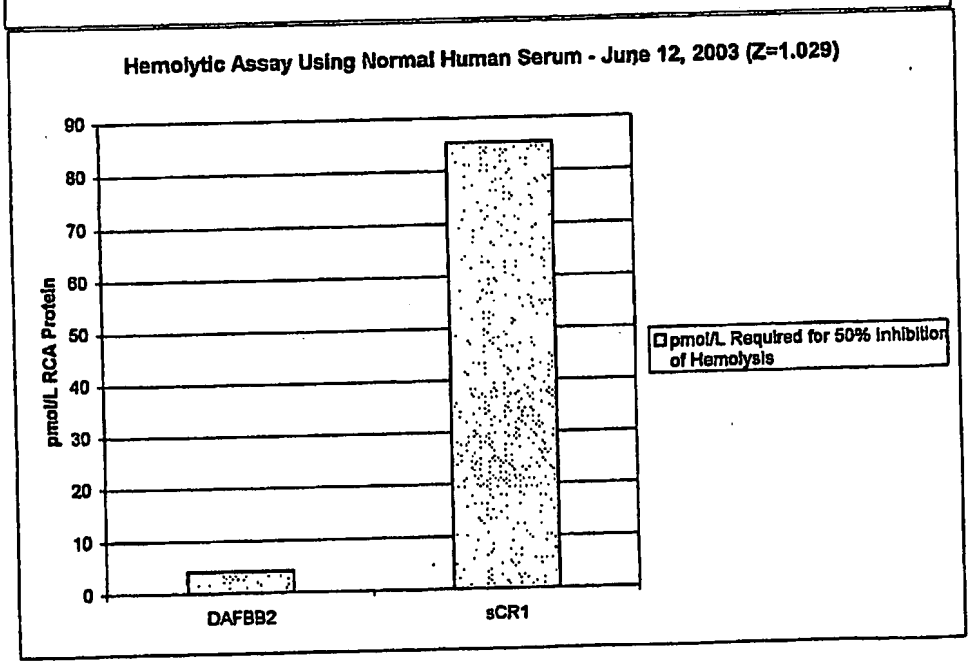
Note: Valine (V) of CR1-IgG4 Linker is the last amino acid of CH1 of IgG4.

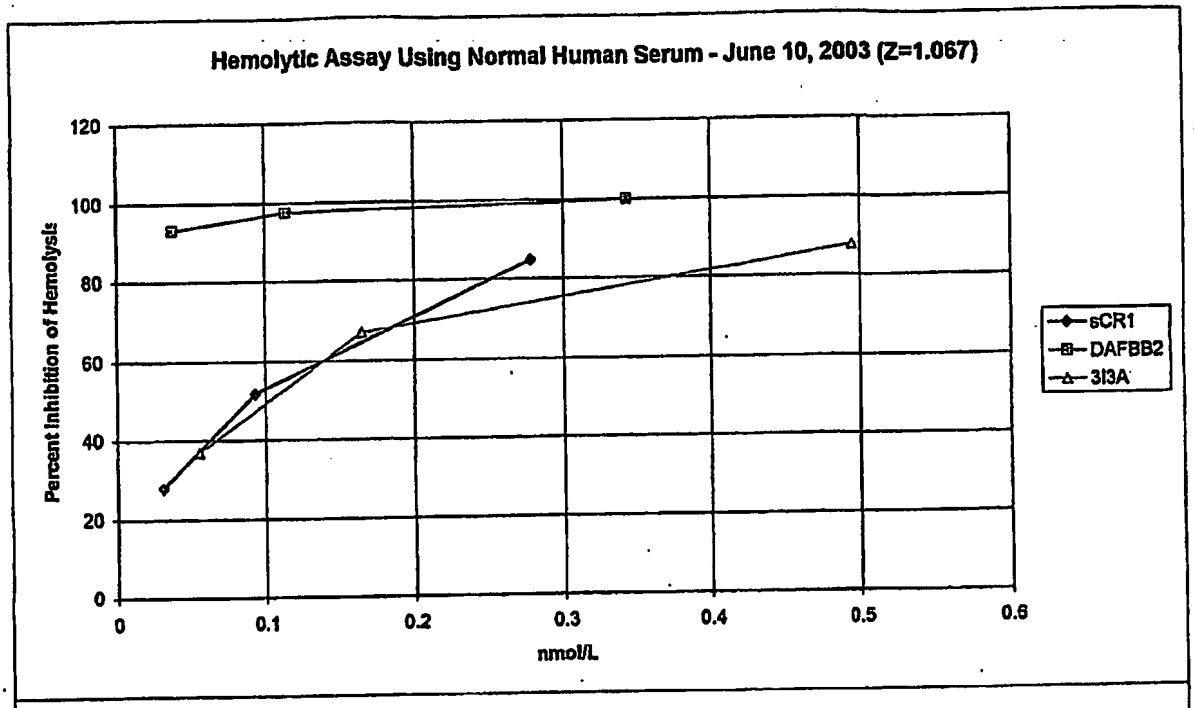
45

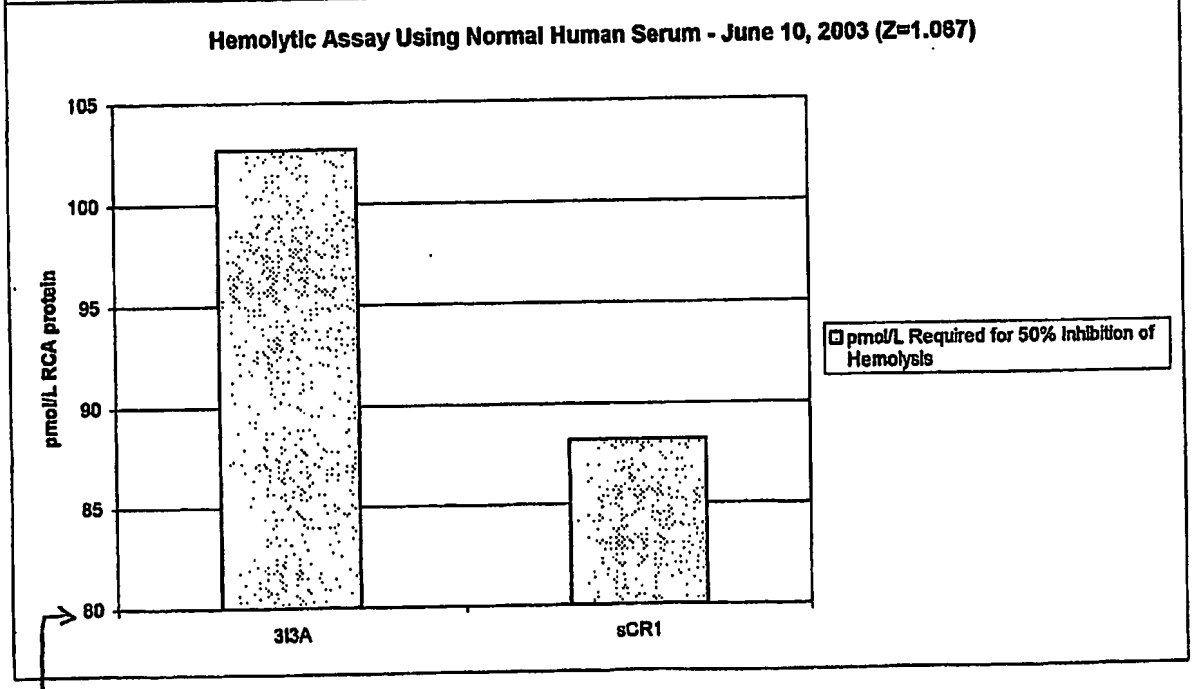
17) Lipid tail for targeting regulatory proteins to tissues.



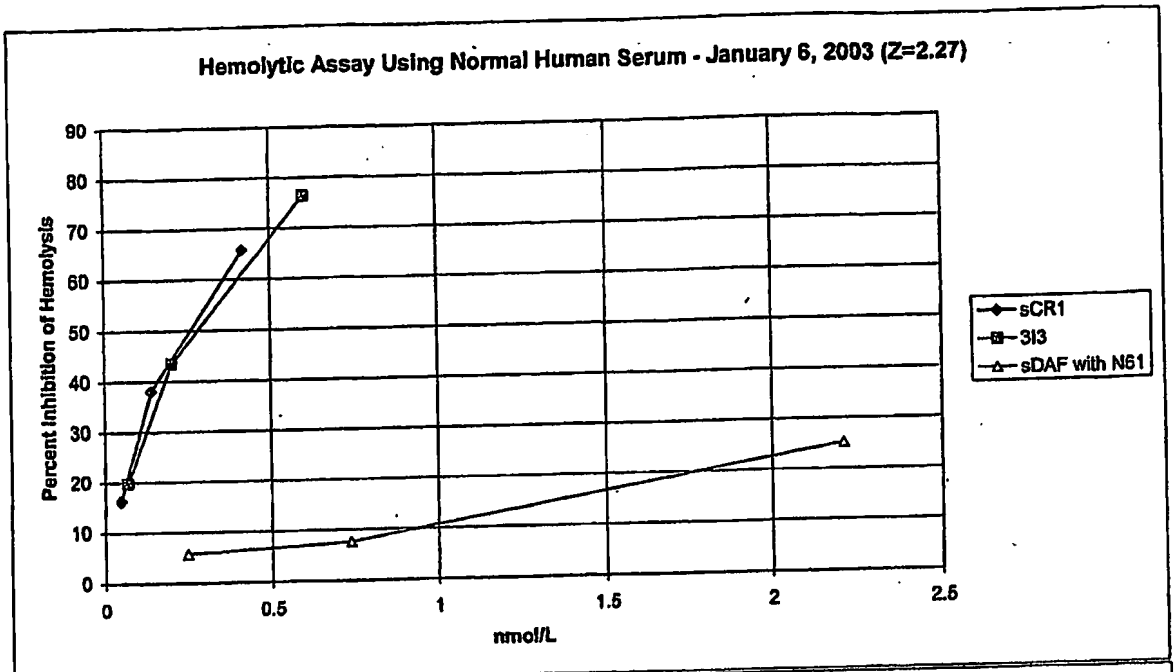


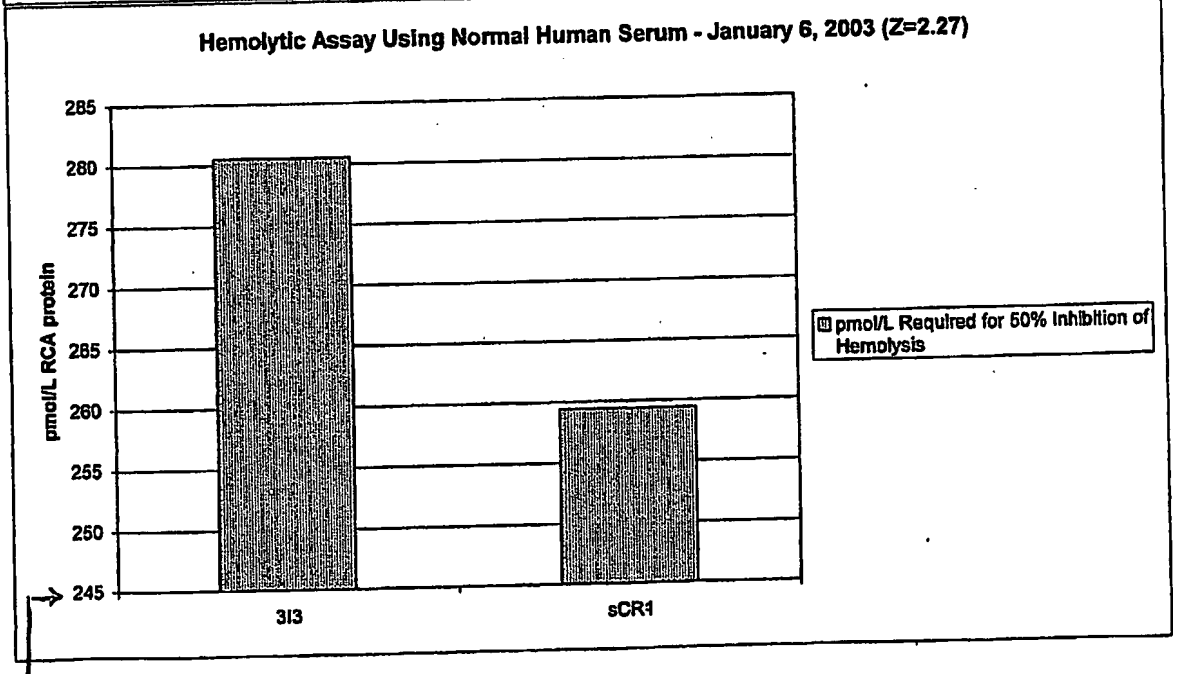



Hemolytic Assay Using Normal Human Serum - June 10, 2003 (Z=1.067)
Percent Inhibition of Hemolysis
120
100
80
60
40
20
0
0
0.1
0.2
0.3
0.4
0.5
0.6
nmol/L
sCR1
DAFBB2
3I3A


Hemolytic Assay Using Normal Human Serum - June 10, 2003 (Z=1.067)
pmol/L RCA protein
105
100
95
90
85
80
3I3A
sCR1
pmol/L Required for 50% Inhibition of Hemolysis

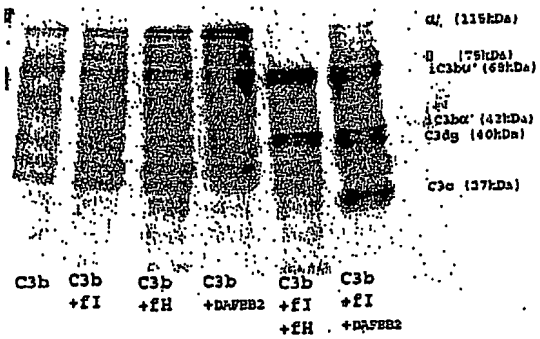
Note Start ≠ 0





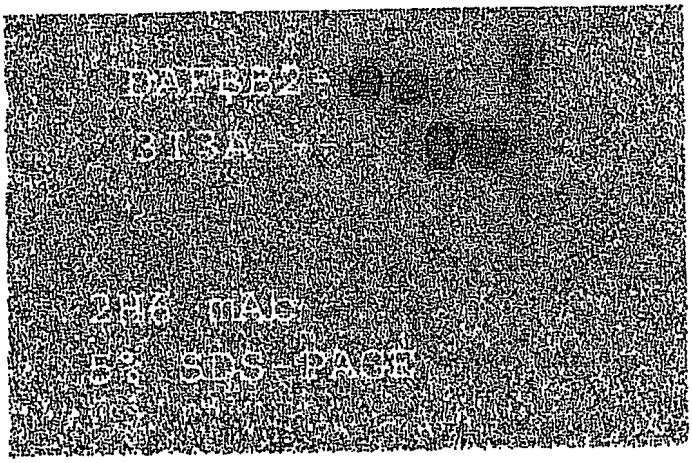
Note Start ≠ 0

Cofactor Experiment (April 22, 2003)



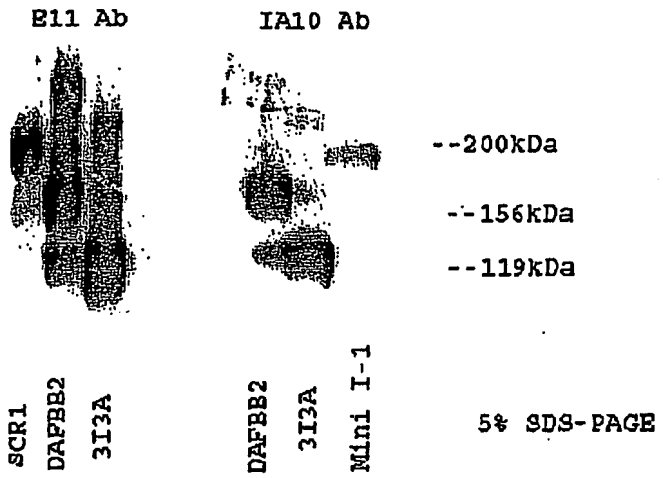
Goat α-human C3 Antibody (ART)
7.5% SDS-PAGE

Western Blot (March 26, 2003)



Note: 2H6 is a monoclonal antibody against DAF CCP 4

Western Blot (July 20, 2003)



Note: E11 is a monoclonal antibody against CR1.
IA10 is a monoclonal antibody against DAF CCP1.